Di Paolo, E. A. (2010). Living Technology. In *Living Technology: 5 Questions*, Mark Bedau, Pelle Guldborg Hansen, Emily Parke, Steen Rasmussen (Eds), Automatic Press/VIP, pp. 67 - 76.

Living Technology

Ezequiel Di Paolo

(1) In what sense do you find it meaningful to talk about "living technology?"

There is a beautiful tension in terms like 'living technology'. Historically, we have tended to contrast the world of machines and the world of living creatures, often seeing them as separate, or even opposed. At the same time we have repeatedly borrowed concepts and metaphors from one world to try to better understand the other. At least since Descartes, animal bodies have been described in mechanistic terms. The metaphors may have changed but this trend has only increased with technological sophistication. We had mechanical, hydraulic and thermodynamic metaphors to depict organic bodies and today we have distributed computer networks as models for complex biochemical regulation, immune reactions, and neural activity. Similarly, we have understood technology in animistic terms, often only heuristically but sometimes with serious consequences. As extensions to the domain of human action, machines are often inspired by natural examples. They can sometimes be regarded as temperamental when they seem keen on pursuing purposes other than our own, making us treat some machines as living beings. The very drive toward automation informs us about the desire for technology to approach the organic and human world as its pinnacle of sophistication and achievement.

These two worlds, that of living organisms and that of human-made artefacts, are also at odds with each other, and herein lies the interesting tension in terms like 'living technology', or similar ones, like 'artificial intelligence', 'artificial life', or 'synthetic biology'. Using machine metaphors to explain life is a useful, but ultimately limited tool for science. And seeking inspiration in organisms to perfect technology sometimes reveals serious contradictions between natural living processes and engineering methodologies. Fundamental disparities between life and (present day) technology become apparent when we consider two essential contrasts between organisms and machines.

Firstly, organisms literally construct each of their components and themselves as integrated wholes in a constant, far from thermodynamic equilibrium flux of matter and energy – this could indeed be taken to be the defining property of life. Anything we consider stable in an organism, whether we are talking about cells, tissues, organs, anatomy, behaviour, knowledge, or social relations, is so only in a precarious sense; things are stable as long as the networked processes of molecular transformation, and several others, keep working. All organic structures are "stable" in a dynamic, non-equilibrium sense, or even unstable but kept within viable bounds, and, as far as we can tell, they all decay eventually. Crucially, this dynamic stability does not rely on material permanence, as is typically the case for the components that make up a machine. In general, each component of a machine remains the same, not by virtue of complex exchanges with other components, but by virtue of its material inertness. We explicitly want machine components to be this way. The resilience of a machine lies largely, though not exclusively, in the resilience of its parts. Machine parts remain machine parts even outside the machine and this cannot strictly be said of organic components. However, far from being a disadvantage, the fact that organic structures are precarious is the root of a different kind of distributed resilience: The autonomy and inherent restlessness of organisms, as well as their flexibility and plasticity, are all properties that we find very hard (nearly impossible) to replicate artificially – all properties, however, that technology envies.

A second and related difference between organisms and machines is that the latter are, without exception, made for satisfying some external goal, while we cannot say the same of organisms. Instead they have an inherently internal purpose: their own continuation as living beings. It makes sense to speak of living organisms as fulfilling an external purpose only when they are used to satisfy human ends (e.g., transport, labour, protection, environmental engineering), in other words, when they are used as machines. Interestingly, our relation to animals becomes less obviously goal-oriented as we interact with them in terms that do not fully deny their autonomy, for instance, in the case of domestic pets. One way to express this contrast is to say that there is a tension between the autonomy of living organisms and the external purposefulness of machines.

Similarly, we find differences moving in the other direction of this relation, from the organic to the technological world. Traditional engineering methods have exploited useful principles of modular design (e.g., divide-and-conquer, combinatorial re-usability, functional compartmentalization). By contrast, "natural design" is often messy, opportunistic, softly modular, functionally distributed, and emergent. It does not have watertight divisions between structures and operates by a (poorly understood) principle of emerging robustness of the whole in the face of unreliable parts, processes, and external influences.

Disciplines like evolvable hardware (i.e., the use of Darwinian principles to evolve functionality in reconfigurable electronic circuits) provide a striking illustration of this difference in how design is approached by nature and by human engineering. In the 1990s, one of the pioneers of this field, Adrian Thompson, from the University of Sussex in the United Kingdom, found something intriguing when trying to evolve electronic chips to do some simple discrimination between different inputs. After many generations, the chips were working as intended, but on studying their evolved structure, Thompson encountered a puzzle: There were components of the chip that were not connected to anything, and yet clamping these components would make the chip not work correctly. How was this possible? To his surprise, Thompson found that these "disconnected" components were in fact interacting through electromagnetic coupling with the rest of the circuit and modulating the processing of input into output in an analogous way. Working directly on the raw material medium allows artificial evolution to bypass constraints that originate in habitual engineering practices, like in this case, that of restricting interaction between chip components to digital information processing and logical gating. Nature, in contrast, doesn't design with a drawing board.

I think it is this double tension that makes 'living technology' an exciting umbrella term. A popular myth about science is that formal deduction is its central driving tool, that scientists extrapolate what the established laws predict and perform experiments to verify them. In fact the driving forces of science are tensions and paradoxes like the ones between organisms and machines. They act as "epistemic gradients" that seek a resolution and fuel creative ideas. This, in my opinion, is what is being denoted by terms like 'living technology'.

(2) How does your research relate to living technology, and why were you initially drawn to do this work?

My work explores the continuities between life and mind. I'm interested in understanding the genuine autonomy of living systems, what makes them active and animated, and what makes them build and follow their own purposes and norms and connect to other life. I see life as already mindful in a general sense, even in "simple" cases like bacteria. I'm also interested in understanding the fundamental organizational principles of life, mind and social interaction so that they could eventually be implemented in synthetic systems. So far, the majority of work in cognitive science, artificial intelligence, robotics and artificial life is still framed in traditional, Cartesian assumptions about the mind being something like the problem solving done by a computer inside the head. Such problems are given to the cognitive system externally (by the environment or others) or artificially imposed by the designer (say, a computational module instantiating some form of value system; a box labelled "Motivation" for instance). These designs are very unlike organisms. The latter instantiate an emergent systemic level and, at this level, they actively distinguish themselves as wholes in relation to their environment. It is only at this level, and not at others below it, like that of internal mechanisms, where the language of cognition applies – terms like 'motivation', 'affection', 'decision making', 'behaviour', 'emotion', 'memory', 'learning', 'experience', and so on are valid only at the level of the whole organism. It is simply a category mistake to think that these things happen inside the head, for instance in the brain.

This category mistake, however, is pervasive and is partly due to the lack of a workable, scientific concept of the organism. Biology, psychology, and neuroscience are more comfortable at levels above or below the individual organism, explaining physiological facts in mechanistic terms and general normativity (why organisms do what they do) in evolutionary terms. This is understandable because we tend to use machines as our best metaphor for studying organisms. We approach these questions from a human engineering perspective, but as I said above, organic design works using different, often counterintuitive, principles. My interest in evolutionary robotics, dynamical systems modelling, embodied, enactive cognition and phenomenology is both a desire to explore these natural principles of organization and at the same time to shake off accumulated preconceptions; a quasi-systematic form of un-learning.

Having come from a background of physics and nuclear engineering, then moving on to evolutionary robotics and biological modelling, I have always been interested in complexity and attracted to that region of problems that are nearly impossible to solve. And the keyword here is of course 'nearly'. My personal preference is to find a proper balance in a research problem. It's boring to work in areas where you feel that a lot has already been done, where there are well-established paradigms and methods that work incrementally even if strong, difficult questions still remain open (e.g., some areas of physics). Precisely because the remaining questions tend to locate themselves at the other extreme of difficulty, I tend to find them boring too. This is the extreme of challenges that are so out of reach that we cannot do much about them, we sincerely don't even know how to formulate them (e.g., understanding consciousness) and they demand a radical revision of our vocabularies, concepts and methodologies. The ideal sweet spot is somewhere in the middle, a place requiring creativity and novel methods and where challenging questions grow, but where there is also a chance of succeeding and innovating.

I see disciplines like artificial life, complexity, evolutionary robotics, synthetic biology, etc. also as technologies of thought, tools for clearing up misconceptions and shaking the foundations of habitual ways of thinking. That's why I'm interested in them.

(3) How is living technology related to overlapping or nearby research

areas, such as nanotechnology, molecular biology, cloning and stem cell research, genetic engineering and synthetic biology? How is it related to social and technological systems such as social networks or information networks, such as the World Wide Web, cell phone networks and electronic banking networks?

Ideally, living technology, conceived as synthetic approaches to understanding life and lifelike processes, should always be in a dialogue with other scientific and engineering disciplines with overlapping subject matters. A proper dialogue should of course go both ways and for mutual benefit. Sometimes this is not so easy; like most interdisciplinary efforts, it requires patience and good communication to overcome terminological, methodological and even community barriers. But when it works, it works well. The best model, of course, is that of a real collaboration between people at the core of each discipline. In this way, living technology can not only draw inspiration from other disciplines, mostly biological ones, but also inform them by providing new tools for knowledge, novel hypotheses and new techniques.

Artificial life is actually a form of theoretical biology. However, theoretical biologists in the past have resisted taking artificial life seriously (sometimes with justification). Still, over the past decade we have seen the gradual but steady adoption by biologists of modelling techniques that were first developed within artificial life. There is now a thriving field of computational biology that owes something to this exchange. So this is possibly another model to follow, apart from direct collaboration: the development of new techniques that override the limitations of more traditional ones.

The same spread of ideas and techniques happens onto wider socio-technological systems, notably in areas of social networks and pervasive computing. Social networking, information networks and portable devices are already changing the way we think, speak and interact with others. These technologies are already getting beneath our skin, like in the case of brain-computer interfaces. However, we mustn't be too surprised about these developments. They are perhaps more overt, but not quite so radically different from learning to do maths, to play a musical instrument or to acquire a new skill or language. The human body is a technological achievement. Cyborg fantasies and realities fall behind the way that humans are always-already artificial creatures. There's nothing natural about the way we walk, talk, dress, manage our physiology, create our own goals, shape our selves as projects, relate to others or even just act and perceive. Human beings are themselves living technology made second nature.

(4) What do you think are the most important open research questions about living technology, and how you think they should be pursued?

I have always found challenging the problems that involve, at any level, some complex *transformation* of organization: from plastic experience-dependent changes in a cognitive system leading to altered behaviours and consequently further plastic changes, to systemic phenomena at evolutionary and ecosystems levels such as niche construction, evolution of group behaviour and the constitution of new units of organization and selection. Something happens that makes things not quite like they were before, some in-homogeneity in time, in the rules of the game that is being played. Several themes like the emergence of new functions, the stability of novel structures, and the influence of the collective and higher levels on micro-dynamics all appear recurrently in problems of this kind. And putting these themes together in a conceptually clear framework or in revealing models is still a serious challenge. It's the kind of problem that before the beginnings of systemic thinking scientists would always try to avoid (because it's not clean enough), and we are still not very comfortable with such problems even now.

In particular, we can think of problems such as how to endow a synthetic agent with genuine emergent autonomy, like that of an animal. I don't see the real challenge as that of building artefacts that would have, say, the intelligence of a cat (not a modest goal at all, by the way). The real challenge is to synthesize artificial systems capable of behaving autonomously like a cat, i.e., of establishing what is good and bad for them on their own behalf and acting accordingly, not because we have installed a chip that tells them that food, shelter and companionship are good things. This is a radical challenge implying an understanding of how values emerge in a system as a consequence of its own precarious, and yet dynamically robust constitution, and in interaction with the world and with others. Such "organismic machines" would have a stake in their world encounters, they could not be simply reset, and their experiences would leave formative traces in their organization. What they do, the experiences they undergo, whom they interact with and how, will shape the kind of systems that they become, but they will always have some freedom in shaping themselves. This is in short the challenge: *to build a free artefact*. The big stumbling block between our current situation and such synthetic autonomous systems is our poor understanding of the links between constitutional and interactional domains in living systems – in other words, how what they are both shapes and is shaped by what they do; how the body shapes the mind and the mind shapes the body.

(5) What do you consider to be the most interesting and important human or societal implications of research and development in living technology?

I have said that a real challenge for living technology will be to understand the emergent autonomy of life. However, is autonomy a genuine technological goal? In so far as living technology works also as a tool of knowledge, the answer should definitely be yes. This follows the idea of knowledge by construction. We can be certain to have gained (at least some) knowledge about something once we can build an instance of it. Traditional tools have failed at fully understanding life and the reasons for this may be deeply rooted in the methodological underpinnings of the established scientific method. Formalisms, the core of scientific explanations, have limitations that natural life seems to transcend (cf. the above example from evolvable hardware). We find it hard to pinpoint the formal foundation of concepts like emergence, circular causation, autonomy, agency and cognition. In his *Critique of Judgement*, Kant already appreciated this fact about life: We can recognize it but reason cannot capture its inherent teleology, its defining features. At most, he claimed, we can use our intuitive grasp of organisms as driven by an internal purpose in a regulative manner, to help us guide our research, but we cannot ground these intuitions rationally. In contrast, recent philosophers like Hans Jonas have overturned Kant's conclusion. If a fact of experience, e.g., the intuitions about the interiority of organisms that we use to recognize life as such, is accessible to us thanks to the fact that we are embodied organisms ourselves – and not merely rational beings – then we must not deny such forms of knowledge, but instead we must extend our current methodologies to better account for them. Constructing autonomous, living artefacts is then a genuine scientific goal, precisely because it may reveal to us new forms of understanding life in all its complexity.

Whether autonomy as such is also a goal to be pursued for other reasons, for example, enhancing our lives, is less clear (it is also not so clear exactly whose lives will be enhanced – genetically modified crops can lead to economic dependence, intensive monocropping in third-world countries and the consequent loss of internal markets). Undoubtedly, the marriage of systems engineering and biological complexity may bring innovative solutions to long-standing problems. It will also bring a lot of hype, maybe future financial bubbles and their explosions. Here we can also expect rationality to sometimes fail. The real advances, sometimes modest but still significant, will often be overshadowed by the flashy, headline-making, but ultimately short of groundbreaking claims of the funding and attention seekers. Artificial intelligence gurus have been promising for over 50 years that machines superseding the cognitive powers of their creators will be among us "within the next five years." Likewise, we can expect similar claims about living technology, and we will have to discern the real advances underneath the hype.

But real advances notwithstanding, we can still ask ourselves whether the quest of synthetic life is driven by practical needs (what's the use of a genuinely autonomous planetary explorer if it decides, autonomously, of course, to stay on Earth because it's safer?) or whether something else is at stake. The ethical concerns around current versions of living technologies, apart from the genuine issues and dilemmas that they raise, also point to a different kind of motivation. The most common moral of stories about living technology, from *Frankenstein* to *Bladerunner*, has been that the child-like but powerful autonomy of the creature inevitably turns against the creator. Despite this resonant, guilt-ridden warning, the promise of mastery over nature has been deeply rooted in Western culture since the Enlightenment. And the dream of creating artificial life has been one of its clearest, not to say most obvious, manifestations. This is an inheritance from modernity that science and technology have not shaken off yet. Among the things to be learned from living technology, in its epistemic use as a tool for new kinds of knowledge, we may find a new understanding of integrative, and still productive, fair and egalitarian ways of relating to nature. Not as masters over it, as capitalism would have it, nor as assimilated components of a super-organic whole, as has been the view of both totalitarian regimes and some environmental movements – these views present us with a static, non-transformative concept of humanity and nature. Instead, we may learn, through novel forms of knowledge, to see nature as the Other in a relation of mutual transformation, ongoing and open-ended. Living technology may serve to redefine both nature and our own human selves.