TU Delft

Delft University of Technology

The Questions of Hermeneutic TA

Towards a Toolbox

Sand, M.

DOI
10.4324/9781003322290-5

Publication date
2023

Document Version
Accepted author manuscript

Published in
Hermeneutics, History, and Technology

Citation (APA)
Sand, M. (2023). The Questions of Hermeneutic TA: Towards a Toolbox. In *Hermeneutics, History, and Technology: The Call of the Future* (pp. 73-88). Article 10.4324/9781003322290-4 Routledge - Taylor & Francis Group. https://doi.org/10.4324/9781003322290-5

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

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

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

The questions of hermeneutic TA - Towards a toolbox[1,2]

***

Martin Sand

***

1) Hermeneutic TA and its objects

With the concept of hermeneutic Technology Assessment (TA), Armin Grunwald proposes a third mode of assessing emerging technologies (Grunwald, 2014b), aside from a first mode that is predictive and quantitative, and a second mode that can be participatory, qualitative and which considers a "corridor of sensible futures." Before I will suggest some programmatic additions to this third mode of orientation in the form of a list of research questions, I will start by briefly examining hermeneutic TA's methodological foundations.

As the term "hermeneutics" in this composition indicates, hermeneutic TA is concerned with understanding. This immediately raises the question, what is the object of understanding in such approach? What shall we understand as technology assessors? Emerging technologies? Strictly speaking, there is no such object as "the emerging technology" that could be assessed. The predicate "emerging" signifies that the technology is incomplete, in-flux, in the process of being development and yet unfinished.[3] Depending on the stage of emergence, these technologies might exist as mere technical drawings, as circumscriptions in funding proposals, as buzzwords in advertisement headlines, or as prototypes, in which case, they would have found their clearest material manifestation (see also Dickel in the present volume). Clearly, however, even in such a stage as prototypes, assessments of the emerging technologies are highly fragile and uncertain: Relevant social and ethical consequences that arise only through

[1] This manuscript is a pre-print version of an article published as Chapter 4 in the book "Hermeneutics, History and Technology: The Call of the Future", ed. by Armin Grunwald, Alfred Nordmann and Martin Sand, Routledge 2023 (https://www.taylorfrancis.com/books/edit/10.4324/9781003322290/hermeneutics-history-technology-armin-grunwald-alfred-nordmann-martin-sand). Please cite this original source and do not share this manuscript.

[2] The following ideas have been presented during a virtual workshop of the HYBRIDA project to which I was kindly invited by Henrik Vogt and Jan Helge Solbakk from the Centre for Medical Ethics at the University of Oslo. I am grateful for the valuable feedback from the project members and participants of this workshop. Here and there, I use "vision assessment" and "hermeneutic TA" interchangeably, since I believe that vision assessment too is primarily interested in understanding and that its objects are more than just technological visions *sensu stricto*.

[3] It is doubtful, whether a technology is ever finished. Individual artefacts (cars, mobile phones, shoes) become obsolete after some time, and, thus, far no technology has ever been considered perfected. Even if that were so, future technologies create innovative pressures, change the socio-technological systems, in which the "perfect technology" is embedded, possibly undermining workability, and may more generally make us see the technology in a less favourable light. Thus, in some sense, all technology is and remains emerging.

societal use (and potential abuse) of those technologies, or their widespread diffusion (individually induced minor harms that collectively cause collateral damage), remain hard to predict (Collingridge, 1980). Not only for an institution such as TA, which is committed to the task of providing information and evaluative orientation for policy-making (Grunwald, 2019), this is a shattering insight. Also other societal institutions and stakeholders are challenged by this: Surely, few of us want to become part of a social experiment with unpredictable and possibly devastating outcomes from which one cannot withdraw and one has not consented to participate in (van de Poel, 2011).

The release and distribution of emerging technologies without previous assessment, sufficient supervision, regulation or more generally, suitable policies would seem to constitute such irresponsible social experiment. Thus, what else – if not the emerging technology itself – is there to understand and assess to provide guidance for this task? It is the pre-diffused, possibly non-material manifestations, visualizations, articulations and verbalizations of the emerging technology that hermeneutic TA is homing in on, that provide the research material that hermeneutic TA aims to study (Grunwald, 2020).

The first reason, we have encountered for doing this is the inaccessibility and inexistence (or not-yet-existence) of the emerging technology itself and the fact that such manifestations serve as suitable proxies for an early assessment. The second reason for doing so, which is related to the first, is that those early visualizations and verbalization of a technology can be seen – if the semantic stretch is allowed – as a trace, if only a blurry one, to their future (Sand, 2018b).

Usually at the point, when TA recognizes the public impact of a technological future, the future has already gained significant traction in society, attracted public attention and its momentum is further increasing (Nye, 2006). While there is a clear significance in this moment of the societal attributing of meaning to a technological future (Grunwald, 2017; van der Burg, 2014), the pathway to its realization even at this point is not certain or determined.

This observation needs qualification and shall be examined briefly, because it also entails a threat for the hermeneutic approach: Is there not a danger, one might ask, that hermeneutic TA is directing its attention to a technological future in the present that will never materialize providing futile orientation, while missing out on developments that impact technological change much more profoundly? Before even getting started, must hermeneutic TA not implicitly endorse a mode 1 assessment of a certain technological vision, predicting the likelihood of the vision to materialize, to determine whether a mode 3 assessment is actually required and justifiable in the respective case? Without such assessment, one might vainly

provide orientation for technologies that are doomed to fail. One can delineate at least three ways in which the connection between a technological future and its realization can be severed or present us with a misleading starting point for an assessment (keeping in mind that there are numerous others, which could be mentioned here): a) Obviously, some technological futures simply fail to materialize for various social or technical reasons. Thus, no matter how hyped and how much significance is ascribed and public attention granted; not every technological future turns into reality – some don't even get close to that. b) Many social developments might actually have a bigger impact on technological change than the debated and hyped technological futures on which hermeneutic TA sets its eyes,[4] as I have outline (see (Sand, 2019).[5] Consequently, shouldn't those developments receive more attention in a hermeneutic TA approach. c) Lastly, some technologies materialize without having ever been envisioned or planned in a forward-looking manner. They come out of the blue and appear suddenly and quasi-accidental. Hermeneutic TA will be caught as much off guard as policy makers and other stakeholders. Taken together, do these possibilities not seriously threaten the meaningfulness of focusing on technological futures to provide orientation?

Having introduced these distinctions, we are now equipped to respond to each of these possibilities and defend hermeneutic TA against the charge of futility and stabbing in the dark. A focus on technological futures to provide normative orientation can be justified despite those possibilities. Focusing on technological futures that are publicly ascribed with meaning and that receive attention, presents us with the biggest chance to yield relevant knowledge to orient policy-making.

Option c), while being clearly possible, is a rare occasion. In every stage of technological change, human decisions are being made (Johnson, 2015). Technologies never assail societies without being frequently pushed and steered by various human decisions. In those rare occasions, where such process is superfast and almost sudden, no evaluative discipline or method can cope. A technology that appears suddenly, or by chance, escapes any sort of anticipatory assessment. Therefore, this possibility does not undermine hermeneutic TA, it presents, in fact, a challenge for all sorts of preemptive assessments. A general instantiation of the precautionary principle as a meta-principle to regulate socio-technical change and slow

[4]

[5] In the problem of choosing, on which of the many present technological futures one ought to focus, the epistemic concern that forms one horn of Collingridge's dilemma partially returns and cannot fully be dissolved by the hermeneutic shift in perspective. If we assess technological future A for its present social significance instead of the less prominent present technological future B, we implicitly presume that social significance of a technological future is relevant for the development of a future technology. This is often, but certainly not universally true. This remains an inevitable epistemic caveat.

such processes down, is the only and most controversial response to this, which exceeds, however, any particular process of concomitant technology assessment and must be based on a more general view regarding the acceptable risks that can be imposed on society without knowing more about the risks of a particular technology. Regarding option b), hermeneutic TA, I believe, has to acknowledge on the one hand to a certain degree that it has, despite its transdisciplinary aspirations, certain boundaries: Its research scope cannot cover all possible sorts of (global) social dynamics that might have an impact on technological change. However, there is no reason that hermeneutic TA could not also try to open itself up to accommodate and look more broadly into those futures that initially appear to lack technological content and that seem mere "mundane" given their potential to nevertheless redirect socio-technical change (Sand, 2019). Such opening up stands in no contradiction with any of its other methodological convictions. Lastly, there is the possibility a) that technological futures are being assessed without materializing. The orientation and normative guidance that has been provided, appears to be futile and resources have been wasted. That view, however, rests on a misunderstanding: First, hermeneutic TA's intent of gaining knowledge clearly aims at – but also goes beyond providing orientation for policy-making. For instance, if we better understand the human condition based on an assessment of a technological future even if that future never materializes, we have already gained significantly. Second, in some instances, putting a halt to the realization of a technological development is not a lamentable development that indicates a waste of TA resources, but is in fact exactly what hermeneutic TA after assessing the technology has recommended. Lastly, the concern presented here has its analogue in the field of statistics as the problem of sensitivity rates: Without being able to expand on this in more detail, it can be said that it is *prima facie* better to assess more technological futures that never materialize than to not assess technological futures that do materialize. Hence, oversensitivity regarding technological futures seems a justifiable stance.

So far, we have established that hermeneutic TA is concerned with understanding the emerging technologies in their present guise as futuristic narratives given the absence of material objects (Grunwald, 2014a). But what does such understanding entail? Which question do we have to ask, to gain a better understanding of future technologies in their manifold present representations? In the following, I will respond to this concern by discussing two examples and propose a number of research questions that might help to better understand the gravity, value and impact of technological futures.

2) Autonomous Weapon Systems, Nanotechnology and Responsibility

How can an understanding of technological futures in their present be achieved? How does hermeneutic TA approach technological futures in the present? In order to understand, how technological futures manifest themselves in the present, we should have a look at two examples a classic and a more recent one.

Clearly, Eric Drexler's "Engines of Creation" is one of the most famous examples of a comprehensive technological vision that initiated a large public debate, before there was a tangible, material technology on the nanoscale. The book written in 1986 proclaims that Nanotechnology "will bring changes as profound as the industrial revolution, antibiotics, and nuclear weapons all in one." (Drexler, 1986) In the book, Drexler envisions life-prolonging molecular machines that would journey through the human body and repair broken, aging cells infinitely, and he imagines nano-factories that absorb carbondioxide out of the air and transform it into oxygen, thus creating an artificial photosynthesis cycle. Nano-assemblers or nanobots could be the smallest and most versatile building blocks in our economy, creating all other machines that are used by humans. It would be an exaggeration to say that the "engines of creation" and Eric Drexler alone initiated the emergence of Nanotechnology. There were other people, too (Smalley, Roco and Bainbridge etc) (Rip & Voß, 2013). But the vision provided a reference point for the community and a medium to reflect the potential of the technology and how to pursue it (Amato, 1991; McCray, 2013). It is fair to say, that without Drexler's vision and effort as an enactor, nano might not have gained the momentum that it got or at least only much later.

In the following second example, too, a scientist has delivered a narrative of a possible future; not to initiate but to halt or at least severely constrain a technological development through regulation. Stuart Russell writes in his book "Human Compatible" that "Autonomous Weapon Systems have been described with good reasons as the third revolution in warfare, after gunpowder and nuclear weapons." (Russell, 2019, p. 110) A page later, while arguing that so called killer robots are often misrepresented in science-fiction films (e.g. Terminator), he says that these weapon systems are actually not science-fiction: They already exist. Surely, as an ordinary reader one might think that Russell, who is a renowned computer scientist at University of California, Berkeley, must know what he is talking about. As an example, he mentions the Harop drone from the Israelian Aerospace Industries, and the "possible design of an autonomous weapon with a small explosive projectile" that is presented in the "Slaughterbots video". On page 111, we find a screenshot from this video that shows a hand carrying a mini-drone.

These claims deserve critical scrutiny: First of all, it is important to point out that the Israelian HAROP drone is in fact „a full man-in-the-loop operation" weapon and not an autonomous drone.[6] Second, what Russell does not mention here or anywhere else in the book is that the Slaughterbot-video, to which he is referring which has been uploaded on Youtube in 2017[7] and to which he refers to underscore the claim that lethal autonomous weapon systems already exist, is in fact a fictional product that he himself created supported by Elon Musks' "Future of Life Institute." The video shows miniature drones acting as swarms, targeting individuals with high precision and killing them effectively with little projectiles (Scharre, 2017). The video has been clicked 4 Mio. times since it was released and has been widely received (Crowder, 2017; Pasquale, 2020). CNNTech has cut the footage and uploaded a shortened version of the clip on their own Youtube channel.[8] In his book, Russell goes on to argue that the "Swiss Defense Department has already built and tested a real Slaughterbot and found that, as expected, the technology is both feasible and lethal" (p. 111).

This is quite an exaggeration: What the Swiss Defence Department really did was – in fact, inspired by Russell's own Youtube video – reenacting a scene from the Youtube clip by placing a 3-gram mini explosive on a simulated head, which was effectively a gelatine-filled polyurethane sphere. They found that the mini-explosive's "detonation caused severe damage to the object."[9] The guiding question of this singular experiment was to test "the extent of the injuries that a miniature shape charge can cause." If we shall, for reasons of charitability, let this experiment indeed count as a validation of the power of mini-explosives, this still puts only a sub-claim of the video to the test. The video proposes a host of other possibilities that – only if taken together – constitute the threat of autonomous mini-drones as presented. I will return to some of these possibilities and implicit presuppositions later. Testing whether the video's larger narrative is actually feasible, would require a very different approach. In the experiment of the Swiss Defence Department, there were no flying objects involved, let alone autonomous one's that act in swarms. It is, therefore, highly inaccurate to conclude from this that autonomous drones that operate in swarm-like fashion, that can accurately identify individual targets and kill them with high precision is a real possibility and it is blatantly false to state that real Slaughterbots exist. Russell has created a narrative and an own hype around a future technology that does not exist yet – and he continues to perpetuate this hype with inaccuracies

[6] https://www.iai.co.il/p/harop (last accessed: 15.07.2021)
[7] https://www.youtube.com/watch?v=9CO6M2HsoIA (accessed, 16.07.2021)
[8] https://www.youtube.com/watch?v=fWzXjn5hHjs (accessed, 15.07.2021)
[9] https://www.ar.admin.ch/en/armasuisse-wissenschaft-und-technologie-w-t/home.detail.news.html/ar-internet/news-2018/news-w-t/lethalmicrodrones.html (accessed, 16.07.2021)

that he himself inserted into the debate. Both the Swiss Defense Department and Russell, pursue their own strategies entering this debate. The Swiss Defense Department states in their press-release (see footnote 8) that the experiment is part of a larger agenda: "armasuisse S+T is actively engaged in anti-drone defence. A multidisciplinary team of sensor and effector specialists are researching defensive options in conjunction with national, international and industry partners." Linking their work explicitly to the public debate about Russell's video, increases awareness for their own work. Russell himself is a sturdy supporter of various NGOs that aim at putting a halt to the development of lethal autonomous weapons. Unsurprisingly, the subtitle of the Slaughterbot video reads: "If this isn't what you want, please take action at http://autonomousweapons.org/."

That this sort of activity and inaccuracy, when it comes to such public relevant issues, causes various, pressing questions about scientific integrity and responsibility is obvious. Questions regarding responsibility can be raised more generally in relation to such sorts of "visioneering" activities (Simakova & Coenen, 2013). An apt description of the concept "visioneering" referring to a number of other illuminating examples can be found in Patrick McCray's, book "The Visioneers. How a group of elite scientists, pursued space colonies, nanotechnologies and a limitless future". McCray writes:

> „[V]isioneering means developing a broad and comprehensive vision for how the future might be radically changed by technology, doing research and engineering to advance this vision, and promoting one's ideas to the public and policy-makers in the hopes of generating attention and perhaps even realization." (McCray, 2013)

What we have encountered previously are forms of visioneering; in the case of Russell, not with the intent to realize a future, but with the intent to put one to a halt. In order to fully understand the meaning of such technological futures in a hermeneutic fashion, I propose to expand hermeneutic TA to also consider visioneers, their actions and their subsequently their responsibilities (Sand, 2016; Sand & Schneider, 2017). This is important, first, because utilizing unfair practices (such as spreading half-truths or exaggerations) might pose a challenge for a more rational debate about the future and the desirability of the envisioned technologies. Calling out such threats to rational debate is part of the normative orientation that hermeneutic TA can provide to improve policy-making regarding emerging technologies. Second, since such behavior could also constitute a normative reason not to trust the expertise of such agents in the future; this remark too helps to orient the debate in suggesting, whose views are creditable and reliable and whose aren't. I will return to visioneering assessment in the following section, when we have a clearer view on the idea of hermeneutic understanding. To summarize, it is

obvious that those futuristic narratives and activities provide an interesting entry point for technology assessment at a very early stage:

> "The factual importance and power of futuristic visions in the governance of science and in public debates are a strong argument in favour of the necessity of providing early public and policy advice in the [new and emerging science and technology] fields [...]. Policy makers and society should know more about these positive or negative visions and their background. They should understand what is going on scientifically and technologically, what is or might be at stake for future developments, […], and who might be affected […]. In summary, this needs uncovering which meaning, values, and interests are hidden in the techno-futures being communicated. Thus, gaining a comprehensive understanding of the meaning of [these developments]." (Grunwald 2014a, 279)

As argued at the outset, factual importance of the vision in present discourses does not ensure the future flourishing of the envisioned technology. It means, however, that hermeneutic TA has an identifiable entry point for assessment that could at least reveal knowledge about the human condition or present interests of relevant stakeholders and providing orientation in cases, where a technological development fails, is still more desirable than missing out on providing orientation for those that succeed.

3) Questions to be asked – Understanding in hermeneutic TA

Having introduced some examples of technological visions, we should now return to the question of what it means to understand technological futures in their present guise as visions and narratives. We saw that in the absence of a material object or prototype, we are in difficulties anticipating the future impact of an emerging technologies. However, that does not mean that we are incapable of extracting other normatively relevant information that might provide orientation and increase self-reflectivity: The visions that have just been presented contain valuable raw material that provide an interesting entry point for an assessment before the materialization and diffusion of a technology. In the most common accounts of hermeneutics, understanding means grasping the meaning of something in its entirety and its embeddedness in a certain context by studying its distinct parts and their relation to another. Understanding a text in this manner could comprise of an interpretation of the biography of the author, her motives and intentions, an analysis of the historical context in which the text was written and of the various parts of the text, the actions and characters of its protagonists and so on. Thereby, a whole of the meaning of the text is being created that exceeds an understanding of its various, distinct parts. According to Jens Zimmermann, hermeneutic

> "understanding is the interpretive act of integrating particular things such as words, signs, and events into a meaningful whole. We understand an object, word, or fact when it makes sense without our own life context and thus speaks to us meaningfully. […] We have not understood a poem, for

> example, when we can merely repeat the words by heart; rather we demonstrate understanding when we intone the words meaningfully and are able to express the poem's ideas in our own words." (Zimmermann, 2015, p. 7)

Thus, as Armin Grunwald suggests, to understand visions and technological futures as a whole, their meaning and place in current debates, requires asking not just a single question but a range of different questions and to investigate their interrelation. The following list has been proposed in his *Technology Assessment in Practice and Theory* (Grunwald, 2018, p. 122):

- What do current narratives of the future signify for the relation of humanity and technology, for humanity and nature? How do they alter or transform these relations and what is regarded as being at stake in ethical, cultural, and social terms?
- Which actors are engaged in the future-oriented debates on new technology in the respective field? What are their motivations, diagnoses and observations, hopes and fears, values and interests, etc.?
- How are narratives of the future communicated and discussed among experts and the public? What roles do they play in the technological debates or neighboring debates about future issues of our time? (Visioneering Assessment)
- What forms of communication and linguistic means are used and for what purposes? What extra-linguistic tools (e.g., movies, works of art, pictures) play a role in this context, and what does their use reveal?
- Is there some hermeneutic significance in the mode 1 or mode 2 orientations mentioned above? Are underlying messages or *Zeitgeist* issues part of the technology trends, scenarios, time lines, and roadmaps presented?

With the intent to substantiate hermeneutic TA as a toolkit that is fruitfully applicable to various technological futures, I will in the following expand on some of these questions and notions and provide some suggestions for extending this list of questions that can be asked from a hermeneutic TA perspective, while explicating their relevance with reference to the previously introduced cases. The interesting aspect about the following set of questions is that they do not aim to predict the future of a certain technology or assess their feasibility. They side-step the epistemic problem that forms one horn of Collingridge's dilemma as introduced at the outset. However, they can nevertheless be extremely helpful in providing normative orientation. We can cluster those questions roughly in broader/anthropological questions and narrower/more specific questions that are not equally applicable to all technological futures.

Broader questions intend to understand not necessarily a particular technological future, but the very fact that technological futures are indeed at all promoted and negotiated in the public realm and that meaning is attributed to them. All technological futures entail certain ideas about the relationship between technology and humans and about human nature and its possible technological transformation. Thus, looking more broadly at the societal constellation in which agents narrate technological futures could help us understand if not the human condition as such then at least modern men. Narrower concerns

a) Broader concerns

- An interesting question one of the broader, anthropological type that isn't trivial at all is: Why are those visions at all projecting and narrating about a **future** technology? Why do they look into the future? One might wonder: Well, why not? When looking further back into human history, we realize that this future orientation is a rather modern phenomenon. It has not always been the case that imaginations and justifications of better societies were located in the future: The early modern utopias of Thomas More and Francis Bacons "New Atlantis" presented visions of different societies that existed on earth at the same time but at another place (Bruce, 2008; Kumar, 1991, p. 57). As Krishan Kumar writes: "The early modern utopia, confronted with a still largely unknown earth, was content to find the good society on isolated islands or strange continents." (p. 57) Even earlier, some cultures looked backwards in time to seek orientation – think about Plato's references to the myth of Atlantis and the vision of Paradise in the old testament (Kumar, 1991). It is only through Millenarist movements and technological change beginning in the 18th century that people started looking towards the future as a reference point where the just and the good society will be realized (Gray, 2008; Nipperdey, 1975; Segal, 2006). Since then, the emphasis on "the future" is ever-increasing and one wonders what this tells us about modernity.
- In which way is the belief in technological salvation metaphysical or religious? David Noble has traced religious motivations behind many techno-scientific endeavours (Noble, 1999). Maybe there is a more straight-forward way to understand ideas such as "mind-uploading" etc. as forms of modern spiritualism (Coenen, 2006).
- Which anthropological view underlies the vision? Many transhumanists think that technology can provide a solution to a great many grieves and human miseries, including ageing and death, because they have a reductionist or naturalist idea of human nature (Sand & Jongsma, 2016). Similarly, many visions of artificial intelligence imply a rather odd view about human intelligence.

b) Narrower concerns

- What's in a name? Nanotechnology has always been an umbrella term for a host of different scientific methods and approaches and also a wide-range of different applications (Nordmann, 2006; Rip & Voß, 2013). Today, nanomaterials can be found in sunscreen or paint. These different applications raise very different ethical concerns. In the visionary stage, when Drexler wrote about the "Engines of Creation" there was no tangible material object to investigate, which makes it very important to be conceptually precise and dissect what exactly is meant by a term such as nanotechnology. The discussion about artificial intelligence (AI) is affected by this problem, too. AI encompasses a wide range of different technologies including, symbolic AI, machine learning and deep neural networks, each having their narrowly confined realms of application and raise their specific ethical concerns (Mitchell, 2020). While machine learning deals very well with pattern recognition, symbolic AI deals much better with rule-oriented tasks. Human intelligence ingeniously combines both (and probably more that cannot be translated or pressed into any rule-oriented shape such as intuitive and emotional intelligence). Visions often narrate the future of a technological **family**, thereby glossing over the oftentimes important factual and ethical differences between their various **subspecies**.
- What is left out by the vision? This is important in two ways: First, visions always simplify the possible future realities they represent, **and** the pathways that lead to them: Technological change is complex and it is remarkable how often new technologies are promised without considering the social factors that are necessary for them to succeed. Drexler is in his book primarily concerned with technical feasibility. But, creating a working prototype of a technology is something is very different, from developing a technology that is diffused and widely employed in society: If that weren't so, we would already have had videophones widely distributed in the 1980s, when they were already possible and not just now. It is the societal context and the existing socio-technical infrastructure that co-determine, whether a technology can gain traction in society. We must ask from a hermeneutic TA point of view: How much have visioneers thought about the "social feasibility" in their narrative?
- Secondly, when we ask, what is being left out of the vision, we also have to recognize in which way many visions forestall answers to the difficult questions about priority setting. By focussing on a very particular concern, visions often gloss over competing concerns, which might even more reasonably quest for public attention. Patrick McCray's visioneering-book contains the following striking anecdote: In August 1977, space

colonization advocate Gerard O'Neill was invited to present his ideas to a thousand or so enthusiasts at the first so-called Space Day at the Museum of Science and Industry in downtown Los Angeles. While Trekkies, NASA representatives and other stakeholders and space enthusiasts gathered inside the Museum, outside the museum, as McCray notes frivolously, "they could have encountered laid-off workers waving signs proclaiming 'Jobs on Earth, Not in Space' and 'Brown, Hire an Earthling' that expressed more terrestrial concerns." (McCray, 2013, 3–4) So what many visions leave out, are all the competing concerns that might be even more weighty than the one presented.

- What are the presuppositions of the vision – the explicit or implicit values? This third aspect ties in with the second: Visions make (often inexplicitly) presuppositions about how society responds to a certain technology: A commentator in IEEE Spectrum does not deny the technical feasibility of Slaughterbots, but points out that the narrators of this dystopian vision presumed aside from technological feasibility that: a) Governments will mass-produce lethal microdrones to use them as weapons, that b) there are no effective defenses against lethal microdrones, c) governments are incapable of keeping such weapons out of the hands of malignant agents, d) and that those agents are actually capable of launching large-scale coordinated attacks with them (Scharre, 2017).
- Visions usually present themselves as a solution to a problem. We must ask: Is the problem that should be solved technologically a real problem and, if so, who has determined that it is and how has it been determined? If it is a problem, does it require technological fixing? In the UK, one can buy for a mere 9,49 Pounds a motorized ice-cream cone, which electrically turns the ice-cream on top: This fixes the "problem" that you have to turn your hand, while eating ice cream. Is that a real problem that requires technological fixing? Another interesting example might be space tourism, which gained recent attention as Sir Richard Branson travelled as a first tourist with a Virgin Galactic rocket into space, thereby celebrating a stage win in the race for space tourism against Elon Musk and Jeff Bezos (Amos, 2021). Space tourism advocates might claim that there is a desire for space tourism, given that there are already sales for space journeys tickets, which shows that there is a **need** for space travels. As said before, we must critically consider what the value of those desires are, if they indeed exist, compared to other possibly more urgent concerns (such as combatting climate change (Mann, 2010)). But even then, isn't the desire a result of the very promise of space tourism as a nearby future technology (van Lente, 2017)? Doesn't

the vision create the need and, therefore, the problem to which it supposedly responds? Is there, in other words, a real problem?[10]

- We should ask: Why now? Many visions centre around technologies that have prototypical precursors or resembling artefacts that precede them. The debate about cybernetics preceded the one about artificial intelligence and many of the questions that are currently discussed have been discussed in similar shape or form some decades ago. If we understand these precursor, what made them fail or stagnate at a certain point and what made their promises return at another, in short, if we understand the hype cycles around these technologies, we provide context to better assess the significance of those technologies (van Lente, Spitters, & Peine, 2013). We can place the technologies into a larger historical development and see, for instance, whether the visioneer fell prey to an ongoing fashion, overestimating the potential of a certain development. As researchers and technology assessors, we are not free of such influences. Stepping outside and getting a broader historical perspective, might help us to remain self-reflective and critical. If can answer, why previous instantiations of a technology failed or challenges remained unanswered, we might be able to concluded why a currently envisioned technology will suffer the same fate.
- Lastly, visions raise pressing questions regarding the responsibilities of visioneers: Who and in which role do agents promote these narratives (Sand, 2018a)? Do visioneers overstep their professional roles and responsibilities? Stuart Russell, Eric Drexler, Ray Kurzweil and others have become major protagonists to drive socio-technical change: Their credibility in ongoing debates and how they conduct themselves, might serve as a soft indicator when it comes to assessing the objectivity and credibility of the technological futures and their risks.

4) Conclusions and outlook

In the present paper, I have suggested that hermeneutic TA can serve as a proper method to assess a variety of technological futures and thereby reveal insights relevant for policy-making and tech-governance. In a first step, I outlined my understanding of how hermeneutic TA responds to the epistemic horn of Collingridge's dilemma by focussing on an assessment of technological futures in the present. In a second step, I suggested that a hermeneutic understanding of such futures must interpret the various elements of those futures and their interrelation to grasp their meaning in a comprehensive manner. This requires locating them in

[10] The reigning paradigm in technophile circles regarding combatting health issues is to produce more "silver bullet" technologies such as antibiotics and vaccines. Given the lack of success in developing more of those silver bullets, more authors propose a shift towards "gentle medicine" (Sand & Jongsma, 2017; Stegenga, 2018)

historical context, analysing the motifs and responsibilities of their promotors and visioneers, and assessing their claims critically.

A final part of such comprehensive hermeneutic understanding, I believe, requires that hermeneutic TA becomes self-aware and critical of its own social and historical embeddedness as an institution that aspires towards democratic legitimacy and value neutrality. This constitutes probably the most pressing concern for future research on hermeneutic TA. Gadamer writes in "On the problem of self-understanding":

> "[…] as a hermeneutical task, understanding includes a reflective dimension from the very beginning. Understanding is not a mere reproduction of knowledge, that is, it is not a mere act of repeating the same thing. Rather, understanding- is aware of the fact that it is indeed an act of repeating. […] The operation of the understanding requires that the unconscious elements involved in the original act of knowledge be brought to consciousness." (Gadamer, 1977, p. 45)

Thus, properly understood hermeneutic TA must become conscious in which way its own unique societal position and the premisses on which its agenda is build, might impact how we assess the desirability and feasibility of certain technological futures and consider other non-technological alternatives as viable. Can hermeneutic TA, for example, accept a future in which a certain problem that has been widely acknowledged as one, is not resolved as a viable alternative to one in which it has been solved with reasonable success and without imposing risk of equal magnitude (Sand, 2021)? The demand for self-reflexivity of hermeneutic TA must in future research lead to the further expansion and examination of the list of broader questions that has been introduced above. By adopting a self-reflective stance, hermeneutic TA really emancipates itself from its roots in vision assessment.

5) References

Amato, I. (1991). The apostle of nanotechnology. *Science, 254*(5036), 1310–1311. doi:10.1126/science.254.5036.1310

Amos, J. (2021). Sir Richard Branson takes off on 'extraordinary' space flight. *BBC News*. Retrieved from https://www.bbc.com/news/science-environment-57790040

Bruce, S. (Ed.) (2008). *Three early modern utopias: Utopia, New Atlantis and The isle of pine*. Oxford: Oxford University Press.

Coenen, C. (2006). Der posthumanistische Technofuturismus in den Debatten über Nanotechnologie und Converging Technologies. In A. Nordmann (Ed.), *Nanotechnologien im Kontext* (pp. 195–222). Berlin: Akademische Verlagsgesellschaft.

Collingridge, D. (1980). *The Social Control of Technology*. New York: St. Martin's Press.

Crowder, L. (2017, 05.12.2017). "As much death as you want": UC Berkeley's Stuart Russell on "Slaughterbots" *Bulletin of the Atomic Scientists*. Retrieved from https://thebulletin.org/2017/12/as-much-death-as-you-want-uc-berkeleys-stuart-russell-on-slaughterbots/

Drexler, E. K. (1986). *Engines of creation: The coming era of nanotechnology*. New York: Anchor Books.

Gadamer, H.-G. (1977). On the problem of self-understanding. In D. E. Linge (Ed.), *Philosophical Hermeneutics* (pp. 44-58). Berkely: University of California Press.

Gray, J. (2008). *Black Mass - Apocalyptic Religion and the Death of Utopia*. London: Penguin Books.

Grunwald, A. (2014a). The hermeneutic side of responsible research and innovation. *Journal of Responsible Innovation, 1*(3), 274–291. doi:10.1080/23299460.2014.968437

Grunwald, A. (2014b). Modes of orientation provided by futures studies: making sense of diversity and divergence. *European Journal of Futures Research, 2*(1). doi:10.1007/s40309-013-0030-5

Grunwald, A. (2017). Assigning meaning to NEST by technology futures: extended responsibility of technology assessment in RRI. *Journal of Responsible Innovation, 4*(2), 100–117. doi:10.1080/23299460.2017.1360719

Grunwald, A. (2018). *Technology Assessment in Practice and Theory* (1 ed.). London: Routledge.

Grunwald, A. (2019). The inherently democratic nature of technology assessment. *Science and Public Policy, 46*(5), 702-709. doi:10.1093/scipol/scz023

Grunwald, A. (2020). The objects of technology assessment. Hermeneutic extension of consequentialist reasoning. *Journal of Responsible Innovation, 7*(1), 96-112. doi:10.1080/23299460.2019.1647086

Johnson, D. G. (2015). Technology with No Human Responsibility? *Journal of Business Ethics, 127*(4), 707-715. doi:10.1007/s10551-014-2180-1

Kumar, K. (1991). *Utopianism*. Buckingham: Open University Press.

Mann, A. (2010). Space tourism to accelerate climate change. *Nature*. doi:10.1038/news.2010.558

McCray, P. (2013). *The visioneers: How a group of elite scientists pursued space colonies, nanotechnologies, and a limitless future*. Princeton: Princeton University Press.

Mead, G. H. (2002). *The philosophy of the present*. Amherst, N.Y: Prometheus books.

Mitchell, M. (2020). *Artificial Intelligence - A guide for thinking humans*. London: Penguin Random House UK.

Nipperdey, T. (1975). Reformation, Revolution, Utopie: Studien zum 16. Jahrhundert. In *Kleine Vandenhoeck-Reihe* (pp. 146). Göttingen: Vandenhoeck & Ruprecht.

Noble, D. F. (1999). *The Religion of Technology: The Divinity of Man and the Spirit of Invention*. New York: Penguin Books.

Nordmann, A. (Ed.) (2006). *Nanotechnologien im Kontext: Philosophische, ethische und gesellschaftliche Perspektiven*. Berlin: Akademische Verlagsgesellschaft.

Nye, D. E. (2006). *Technology matters: Questions to live with*. Cambridge; Mass: MIT Press.

Pasquale, F. (2020, 15.10.2020). 'Machines set loose to slaughter': the dangerous rise of military AI. *The Guardian*. Retrieved from https://www.theguardian.com/news/2020/oct/15/dangerous-rise-of-military-ai-drone-swarm-autonomous-weapons

Rip, A., & Voß, J.-P. (2013). Umbrella Terms as Mediators in the Governance of emerging Science and Technology. *Science, Technology & Innovation Studies, 9*(2), 39–59.

Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*: Viking: Penguin Random House.

Sand, M. (2016). Responsibility and Visioneering—Opening Pandora's Box. *NanoEthics, 10*(1), 75–86. doi:10.1007/s11569-016-0252-7

Sand, M. (2018a). *Futures, Visions, and Responsibility: An Ethics of Innovation*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Sand, M. (2018b). How the future has a grip on us. *Techné: Research in Philosophy and Technology, 22*(2), 262-282. doi:10.5840/techne201862684

Sand, M. (2019). On "not having a future". *Futures, 107*, 98-106. doi:https://doi.org/10.1016/j.futures.2019.01.002

Sand, M. (2021). RRI und Postwachstumsökonomie als Herausforderungen für die Technikfolgenabschätzung. In S. Böschen, A. Grunwald, B.-J. M. A. Krings, & C. Rösch (Eds.), *Technikfolgenabschätzung: Handbuch für Wissenschaft und Praxis* (1 ed., pp. 403-414). Baden-Baden: Nomos.

Sand, M., & Jongsma, K. R. (2016). Towards an Ageless Society: Assessing a Transhumanist Programme. In E. Domínguez-Rué & L. Nierling (Eds.), *Ageing and Technology* (1 ed., pp. 275–294). Bielefeld: Transcript.

Sand, M., & Jongsma, K. R. (2017). The usual suspects: Why techno-fixing dementia is flawed. *Medicine, Health Care and Philosophy, 20*(1), 119–130. doi:10.1007/s11019-016-9747-9

Sand, M., & Schneider, C. (2017). Visioneering Socio-Technical Innovations: A Missing Piece of the Puzzle. *NanoEthics, 11*(1), 19–29. doi:10.1007/s11569-017-0293-6

Scharre, P. (Producer). (2017, 15.07.2021). Why You Shouldn't Fear "Slaughterbots". *IEEE Spectrum*.

Segal, H. P. (2006). *Technology and Utopia*. Washington D.C.: American Historical Association.

Simakova, E., & Coenen, C. (2013). Visions, Hype, and Expectations: a Place for Responsibility. In R. Owen, J. R. Bessant, & M. Heintz (Eds.), *Responsible Innovation* (pp. 241–266). Chichester, UK: John Wiley & Sons.

Stegenga, J. (2018). *Medical Nihilism*: Oxford University Press.

van de Poel, I. (2011). Nuclear Energy as a Social Experiment. *Ethics, Policy & Environment, 14*(3), 285-290. doi:10.1080/21550085.2011.605855

van der Burg, S. (2014). On the hermeneutic need for future anticipation. *Journal of Responsible Innovation, 1*(1), 99-102. doi:10.1080/23299460.2014.882556

van Lente, H. (2017). *De ruimtetoerist: vragen over technologie en toekomst*. Inaugural Lecture. Maastricht.

van Lente, H., Spitters, C., & Peine, A. (2013). Comparing technological hype cycles: Towards a theory. *Technological Forecasting and Social Change, 80*(8), 1615-1628. doi:10.1016/j.techfore.2012.12.004

Zimmermann, J. (2015). *Hermeneutics: A very short introduction* (1 ed.). Oxford: Oxford University Press.