Ignorance at the Heart of Science?

Incredible Narratives on Brain-Machine Interfaces

Alfred Nordmann

Technische Universität Darmstadt

www.philosophie.tu-darmstadt.de/nordmann

Please note: The following is a penultimate draft for the proceedings of a NanoBioRaise workshop. It does not provide an outline of my project in our ZiF research-group. It refers to that project only in passing (footnote 5). However, this paper does provide a context for my work in the research project and beyond that – it is also a somewhat passionate expression of what might be at stake in an "ethics of knowledge assessment" (Rene von Schomberg) and the consideration of the quality of knowledge produced under "mode 2"-conditions.

\- - - - -

On first sight, the workshop title *Making the Human-Machine Interface More Specific* sounds unnecessarily cumbersome. Why not more directly *The Ethics of Human-Machine Interfaces* or, if this were to be a technical workshop, *A Roadmap for Brain-Machine Interfaces*? The cumbersome title tells a story, of course, and it is the story of needing to know what to believe before one can engage in ethical or societal deliberation. This does not answer the question, however, but raises a new one: Where does this special need-to-know come from, how novel or unique is it, what is different here from the cases of reproductive medicine or global warming?

Normally, as with reproductive medicine or global warming, all one needs is the description of a practice ("a woman's egg is removed and then…") or a newspaper report that CO_2 emissions are a likely cause of sea levels to rise, that climate patterns agree with global warming forecasts, etc. Even if there remain differences of scientific opinion, journalists and citizens can nevertheless settle upon a belief: They can find out what the majority of scientists think, they might consider the apparent political and economic (in)dependence of researchers, some will choose to side sympathetically with the brilliant renegades who hold minority views, others will have a sense of how global warming theories are faring as new evidence comes in, and many appeal to their brute experience of hotter summer, less severe winters, more hurricanes and the like. To be sure, there remains a good deal of uncertainty and it would be helpful if science and politics demonstrated greater unanimity. But though our position of knowledge or ignorance is far from ideal in regard to global warming, we can arrive at knowledge that is good enough to underwrite political efforts and ethical deliberation.

If we want to know about human-machine interfaces, however, matters don't seem to be as easy as all that. We read the newspaper, we look at reports and roadmaps, we attend conferences, consult journals – and still don't know what to believe. The usual strategies for arriving at an informed opinion don't seem to get traction. Before a responsible engagement can begin and before genuine ethical and societal issues can be explored, we therefore need to confront a situation that is perhaps as unusual as it is bewildering: We are asked to believe incredible things, we are offered intellectually engaging and aesthetically appealing stories of technical progress, the boundaries between science and science fiction are blurred, and even as we look to the scientists themselves, we see cautious and daring claims, reluctant and self-declared experts, and the scientific community itself at a loss to assert standards of credibility. Quite properly, therefore, responsible ethical questioning doesn't jump right in to confound this situation but steps back to reflect on ways to identify the issues, "making the human-machine interface more specific."

The following remarks provide a tapestry of impressions that illustrate and explore this bewildering situation. They conclude by offering a document of one attempt to address the predicament and move beyond ethical reflection under conditions of incredibility.

Bewildering Fabrications.

One example shall serve to illustrate that we are asked to believe incredible things, that the semblance of credibility is suggested by intellectually engaging and aesthetically appealing stories, that the boundaries between science and science fiction have become blurred, that among researchers one sees cautious and daring claims, reluctant and self-declared experts, and that finally the scientific community is at a loss to assert or even thematize or debate standards of credibility.

When it comes to the nanosciences and the convergence of technologies enabled by them we are asked to believe, for example, that a new kind of manufacturing will emulate nature by building things atom by atom from the ground up. “Molecular manufacturing” or “molecular fabrication” takes any worthless and abundant material, decomposes it into its atomic components and reprograms it to assemble into any molecular configuration. This is to allow for absolutely waste-free production – an environmentally friendly way to achieve global abundance (Crandall 1996).

We are here asked to believe something incredible, for sure, if only because it so closely resembles the alchemical dream of making gold from base materials (though gold has now been displaced by steaks or a good Bavarian *Nanoschnitzel*). We have been taught to distrust this dream and are astounded when promoters of nanotechnology actually compare it to alchemy. For example, one of the founding documents of the US National Nanotechnology Initiative makes this comparison in the immediate context of a proclamation by Nobel prize winner Roald Hoffman that nanotechnology offers “a way of precise, controlled building, with incidentally, environmental benignness built in by design” (Amato 1999, p. 4).

There is another, even more striking reason, why such claims appear incredible – and this is simply because we are actually supposed to find them so. Nanotechnology is to be marveled at, it seeks to inspire a magical sense of wonder at what science and technology is achieving (Nordmann 2005). A particular rhetorical strategy points at the wonderfully intricate workings of a human cell – a miracle of nature or so it seems – and then goes on to take this example of nature's nanotechnology as proof of concept for molecular manufacturing: Nature's alchemy transmutes base materials into complex structures and uses a genetic program to do so; human nanotechnology simply follows nature's lead.[1]

When we are asked to believe incredible things, help is offered by way of engaging and aesthetically pleasing storylines that blur science and science fiction. Various strategies of extrapolation are used to suggest that molecular manufacturing is historically inevitable – that it is only a next step on the long road of technological progress. We will encounter some of those seductive narratives below. The storyline behind molecular manufacturing is that of the next industrial revolution which is to follow on previous industrial revolutions where all of these revolutions are thought to be enabled, if not immediately unleashed by technological breakthroughs: steam, electricity, molecular assembly. This suggestive narrative is already a kind of fiction[2], but the blurring between science fact and science fiction continues as we look to the scienitific community for help. If the vast majority of scientists refuses to take molecular manufacturing seriously, the public does not usually hear from them: The views of nay-sayers are not particularly interesting and members of a silent majority don't have an incentive to invest time and energy just to "set the record straight." The experts in the limelight of public presentations or media coverage tend to be enthusiasts of some kind or another and there are few tools to distinguish between credible and incredible claims especially when these are mixed up in haphazard ways. The vision of molecular manufacturing, for example, cannot be dismissed as "typically American" nano-hype but can be found between the lines also in European documents [3] or in a television appearance of the

---

[1] The question whether technical possibility includes everything that nature can do (or everything that does not contradict outright the laws of nature) was at the heart of the debate between Richard Smalley and Eric Drexler. Smalley dared what few of his colleagues are willing to do – distinguish between physical possibility and technical possibility (by pointing out, for example, that nature's nanotechnology is wet and that the vision of molecular manufacturing is inextricably rooted in dry mechanisms), Drexler and Smalley (20??). Richard Jones sides with Smalley against Drexler but proposes "soft machines" as a third way (2004). The verdict is not out on Jones's hybrid vision.

[2] For a critical perspective on this historical simplification see Bruland and Mowery (2004).

[3] Two examples: "..." (European Commission 2004, p. 10). „The expected STREPs should foresee interdisciplinary research, that may include modelling, 'scaffolds', growth, positional assembly, self-replication

highy respected German nanoresearcher Wolfgang Heckl (Bayerischer Rundfunk 2003). One might think that the matter is setlled in favor of molecular manufacturing when nobel laureates Gerd Binnig (1989) or Heinz Stoermer (20??) add their endorsements. However, the originator and strongest advocate of this vision is Eric Drexler, a renegade scientist who brought nanotechnology to the fore, who conjured doomsday scenarios and who is now considered by most to be a liability for “real” nanotechnological research. On the one hand, the research community distances itself from Drexler almost unanimously, but on the other hand, some established members of that community appear to embrace or at least playfully entertain his most extreme and supposedly incredible vision, that of molecular manufacturing. Confounding matters even more, some of the “experts” that are quoted in the press and invited to conferences are entirely self-appointed – while they speak as experts on ethical and societal issues (in this area one gets away with self-ascriptions of expertise), they use the formula “if it has social impacts, it must be real” to persuade us not of this or that ethical principle but of molecular nanotechnology. This is how Eric Drexler’s Foresight Institute worked in the early years and this how the so-called Center for Responsible Nanotechnology operates today. In a recent edition of its newsletter Chris Phoenix provided a brief description of molecular manufacturing. It contained a number of steps, none of them achievable by current means, and one more ambitious than the other. This, for example, is only the first item on the list: “Build a system that does a billion chemical reactions, one after the other, on the same molecule, with very high reliability, to make perfect molecular products” (Phoenix 2006).[4] But before he can ask questions about the ethical and societal implications of

---

of nano-entities, their functionalisation, etc. Production issues could also be addressed, in particular in view of realising three-dimensional systems towards the fabrication of more complex structures with great and wide potential. Incremental improvements of existing production lines are excluded” (EC 2003, pp.12f.).

[4] “Here in a nutshell is the molecular manufacturing plan: Build a system that does a billion chemical reactions, one after the other, on the same molecule, with very high reliability, to make perfect molecular products. The system does chemical reactions by holding molecules and moving them into place through a vacuum, to transfer atoms to the product, adding a few atoms at a time to build molecular shapes. Use that system to build nanoscale machine components, and assemble the components into nanoscale machines. Control a bunch of these machines to build more machine components, one deposition at a time; then combine those machine components into large products. This will need huge numbers of machines, arrayed in a factory. Use an

molecular manufacturing, he needs to convince his readers that he is offering a realistic vision. He does this by openly acknowledging the many objections and then proceeds to seemingly address them. However, what looks like a credible refutation of a serious objection may just be a meaningless assertion if not reductio ad absurdum:

> As we will see, nearly every phrase in this description may evoke skepticism from someone; however, all of these objections, and many others, have been addressed.
> […]The following is an imagined conversation between an MM researcher (MMer) and a room full of scientists who are new to the ideas.
> MMer: OK, we're going to build a system that does a billion chemical reactions, one after the other, on the same molecule, with very high reliability.
> Chemist: Wait a minute. 99% is an excellent yield, but 99% times 99% times 99%... a billion times is a big fat ZERO. You would reliably get zero molecules of desired product.
> MMer: A chemist is used to reactions between molecules that bump into each other randomly. In molecular manufacturing, the molecules would be held in place, and only allowed to react at chosen locations. Yield could be many "nines" better than 99%. […]
> As you can see, each objection brought by intuition from within a specific field has an answer that comes from the interdisciplinary approach of molecular manufacturing theory. We are not, of course, asking anyone to take it on faith that molecular manufacturing will work as planned. We are only asking newcomers to the ideas to refrain from snap judgments that it can't work for some apparently obvious reason.
> (Phoenix 2006)

Our bewildering situation is characterized by our apparent inability to rule on such expertise: Chris Phoenix behaves like an expert, after all, by succinctly presenting an advanced technical vision with proper detail and understanding, by citing objections and answering them, by

---

initial small factory to make another bigger factory, repeating enough times to grow to kilogram scale. Use the resulting big factory to make products from downloaded blueprints" (Phoenix 2006).

taking seriously also the societal implications of this vision – so, who is to say and what does it matter that there is still little reason to believe that this vision will ever be realized, let alone within the next couple of decades as prophesied by the Center for Responsible Nanotechnology? And where does this leave public debate and ethical deliberation with their need to distinguish credible from incredible claims?

Conceptual Cluelessness.

So far, this bewildering situation has been described primarily as a predicament of the public vis-a-vis nanotechnological visions such as molecular manufacturing. But there have already been suggestions that the scientific community is also implicated in the dilemma, that it helps produce confusion, that it also finds itself confused and at a loss to assert criteria of credibility: There is ignorance at the heart of science, and that is not the kind of ignorance that will be overcome by a quest for knowledge, but an endemic ignorance that results from inherent difficulties in pursuing certain lines of critical questioning. This is a strong claim, of course, too strong to be defended here in a systematic manner.[5] But as we turn to the issue of brain-machine interfaces, further anecdotal evidence becomes readily available. Consider, for example, a 2002 report that was co-sponsored by the US National Science Foundation and Department of Commerce:

> Despite moments of insight and even genius, the human mind often seems to fall far below its full potential. The level of human thought varies greatly in awareness, efficiency, creativity, and accuracy. [...] All too often we communicate poorly with each other, and groups fail to achieve their desired goals. Our tools are difficult to handle, rather than being natural extensions of our capabilities. In the coming decades, however, converging technologies promise to increase significantly our level of understanding, transform human sensory and physical capabilities, and improve

[5] Some hints will be provided here and a more sustained analysis is in progress.

> interactions between mind and tool, individual and team. [...] Fast, broadband interfaces directly between the human brain and machines will transform work in factories, control automobiles, ensure military superiority, and enable new sports, art forms and modes of interaction between people. [...] Individuals and teams will be able to communicate and cooperate profitably across traditional barriers of culture, language, distance, and professional specialization, thus greatly increasing the effectiveness of groups, organizations, and multinational partnerships. [...] New communication paradigms (brain-to-brain, brain-machine-brain, group) could be realized in 10-20 years. [...] Components of wearable computers could be packaged in scintillating jewelry, automatically communicating thoughts and feelings between people who are metaphorically and electronically "on the same wave length." (Roco and Bainbridge 2002, pp. 4, 16, 18)

Dan Sarewitz took part in early discussions that eventually led to the quoted text. As a social scientist he later reflected on this experience and effectively formulated the notion of ignorance at the heart of science:

> Most of the attendees were highly intelligent white males who worked in the semiconductor industry, at national weapons laboratories or major research universities. At one point, the group got to talking about how we might soon achieve brain-to-brain interfaces that would eliminate misunderstandings among humans. Instead of having to rely on imperfect words, we would be able to directly signal our thoughts with perfect precision. I asked how such enhanced abilities would get around differing values and interests. For instance, how would more direct communication of thought help Israelis and Palestinians better understand one another? Unable to use the ambiguities and subtleties of language to soften the impact of one's raw convictions, might conflict actually be amplified? A person at one of the meetings acknowledged he 'hadn't thought about values', while another suggested that I was being overly negative. [...]

> This sort of conceptual cluelessness is rampant in the world of techno-optimism.
>
> (Sarewitz 2005, quoted in Miller and Wilsdon 2006, pp. 21f.)

How can it be that such conceptual cluelessness is rampant in the world of science and engineering? Only a few factors can be mentioned here, their full significance needs to be explored elsewhere.

While scientists are usually throught of as champions of criticism, the criticisim of theories and hypotheses requires the framework of a discipline, of shared concepts, methodologies, theoretical frameworks. In the world of technoscience, however, the most attractive projects encompass or recruit multiple disciplines – and in order to recruit a multidisciplinary research team, mutual trust and some degree of credulity become virtues and necessities. Since everyone knows just how difficult one's own piece of the work is, one would like to imagine that that of the others is comparatively easy. This is necessary as to convince oneself that real progress towards achievement of the project goal is possible at all.

There are other reasons why many scientists have an accommodating ignorance as a normal part of their daily lives – and these have to do with the instruments they use. Often these instruments are (nearly) as opaque to them as are the insides of the computers to most of their users. Indeed, some of these instruments (computer simulations, in particular) are useful precisely because they work beyond the limits of what the human mind can calculate or know. When such instruments, moreover, suggest immediate visual access to their objects of interests, there is little incentive to ask critical questions about artefacts or about alternative ways of representing the data. Indeed, some instruments render obsolete even the very idea of an observer's perspective, a perspective that reminds the observer of the constructedness of representations and the difficulty of relating one's detached standpoint to a remote world "out there." Instead, for many scientists and their students real and virtual worlds become conflated as they "fly through" a physical system (the human body, a molecule), exploring it as if from within.

The apparent difficulty even by the scientific community to assert and maintain exclusive standards of credibility is aggravated further when scientists who do not approve of hype are encouraged to produce exaggerated claims by funding agencies that do not approve of hype either:

> The primary objective of this thematic area is to promote real industrial breakthroughs, based on scientific and technical excellence. Radical breakthrough can be achieved through two complementary approaches:
>
> - creation of new knowledge;
> - new ways of integrating and exploiting existing and new knowledge.
>
> This requires changes in emphasis in Community research from short to longer term as well as in innovation, which must move from incremental to radical innovation and breakthrough strategies, while emphasising an integrating approach. (EC 2003, p. 11)

Authors of research proposals and their peer reviewers are challenged here to leave aside the ordinary scientific business of incremental innovation and of building step-by-step on previous work. They are asked to go beyond their traditional ways in order to pursue radical innovation and breakthrough strategies. Should we be surprised if they gloss over the difficulties of cognitive and social psychology and linguistics, if they risk “conceptual cluelessness” in regard to these other fields in order to advance visions of progress in their own field of information and communication technologies?[6]

Seductive Narratives.

The draft minutes of an exploratory workshop on potential benefits of brain-computer and brain-machine interfaces (BCI-BMI) offer yet another reason for ignorance or credulity at the heart of science. When ambitious programs have only scant evidence to show for so far, grand

[6] Jutta Weber has shown for the case of robotics that the designers of emotional intelligence will – of course! – look to the simplest, and most naive theory of human emotions. The more our communication of emotional states resembles our relation to pandas, the easiest to construct emotional robots. (And for the purposes of research, this is indeed the ideal point of departure, see Weber ??).

narratives are needed to interpret the available evidence as a sign for much greater things to come:

> Current BCI-BMI applications are one-way communication systems (e.g. spelling devices). Current state-of-the-art allows processing 40 bits/min. This is far too slow for effective communication, but a large improvement with respect to a couple of years ago (only 2 bits/min in 2002). Assuming a similar rate of progress, a ***communication speed similar to natural speech*** might be achieved ***by 2020***. (Workshop 2006)

This short paragraph has a story to tell. So far, tremendous therapeutic progress has been achieved by way of implanted electrodes, patient training and software development to help totally immobilized persons spell out words on a computer screen through efforts of concentration.[7] In 2002, patients were able to transmit 2 bits per minute, four years later this figure is up to 40 bits or five letters per minute. If this rate of progress were to continue indefinitely, we can calculate that by 2020 these patients might be able to communicate to the computer as fast as healthy people speak (and that by 2025 everyone can communicate faster than they have been able to so far[8]). The draft minutes do not claim that this extrapolation will actually hold, nor do they call it into question. Instead, they invoke as a standard of sorts for envisioning the future potential of brain-machine interfaces. And what makes this a standard model is Moore's Law that has fairly impressively plotted exponential growth over the last several decades in computing speed, power, and cost. Moore's Law, in turn, has been extrapolated to hold for all technological developments in particular by Ray Kurzweil (2000 and 2005). Kurzweil's attractive curves of exponential technical growth have impressed scientists and policy makers alike. The draft minutes show that they have become a rather ordinary part of the vocabulary of technology foresight.

---

[7] These and similar advances are a far cry from „controlling a machine by thought alone" as these are often reported in the media.

[8] Incredible as it sounds, this is of course part of the previously cited US-American vision: „Visual communication could complement verbal communication, sometimes replacing spoken language when speed is a priority or enhancing speech when needed to exploit maximum mental capabilities" (Roco and Bainbridge 2002, p. 16).

To be sure, this crude template is not endorsed by historians of technology, it is statistically questionable. Even if all the past data were collected and plotted correctly, it clearly runs into the problem of induction that is as old as David Hume’s *Treatise on Human Understanding* (17??). Even if the development of technology had been accelerating throughout history, will it therefore continue to do so? Rather than ask whether such a "historical law" is even remotely plausible, however, its proponents are challenged to imagine what technologies will deliver these growth rates into the future (Hanson 2006).

The notion of exponential growth makes a claim about the history of technology that is as strong as it is seductive. In this and other ways it bears close resemblance to the previously noted storyline about a succession of technologically induced industrial revolutions. But there are other, seemingly less spectacular ways to read small success stories regarding brain-machine interfaces as signs of a future that will see radically enhanced human beings. John Harris (2006) or Arthur Caplan (2006) make incredibly weak claims that also fail to pass muster with historians of technology. They extrapolate from the past to the future by making all technology look alike, maintaining that all technology serves the creation of tomorrow's people who enhance their capacities.

> That’s what agriculture is. That’s what plumbing is. That’s what clothes are. That’s what transportation systems are. They are all attempts by us to transcend our nature. Do they make us less human? (Caplan 2006, 39)

According to these arguments, we either believe incredible claims regarding the technologically enhanced people of tomorrow or we deny the obvious truth that we have always used technology and often pursued some idea of human betterment.

This brings us finally to a strategy that refrains from grand narratives but forces us to bet on what the future might hold. It does not deny that claims of an expanded brain-machine intelligence are incredible. Adopting a tactic of anti-Darwinian creationists, it simply suggests that in the absence of proof we need to take seriously even incredible alternatives:

> [...] to assume that artificial intelligence is impossible or will take thousands of years to develop seems at least as unwarranted as to make the opposite assumption. At a minimum, we must acknowledge that any scenario about what the world will be like in 2050 that postulates the absence of human-level artificial intelligence is making a big assumption that could well turn out to be false. It is therefore important to consider the alternative possibility: that intelligent machines will be built within 50 years. (Bostrom 2006, p. 41)

Our state of ignorance about the future is here taken to imply equiprobability: If we can't be sure that something is impossible, we need to consider its possibility. Instead of seeking better information and instead of focusing on the consequences of current technical developments, we are asked to consider the ethical and societal consequences of something that we may or may not have good reason to believe.

Shifting Grounds.

This survey of a bewildering situation sets the stage for one more or less desperate attempt to deal with it. Invited to discuss world regional perspectives at a World Forum devoted to “Tomorrow’s People,” I found myself in the position of debating speculative prospects of human enhancement (Nordmann 2006). Since I don’t believe that these prospects are credible in the first place, since other ongoing technological developments are likely to pose a far more profound challenge, and since the conference did not begin by “making brain-machine interfaces and other human enhancements more specific,” it seemed clear to me that even a dissenting viewpoint could only serve to legitimize a debate that lacks legitimacy. I therefore found myself in the impossible, indeed bewildering situation of wishing to point out the irresponsibility of the type of questioning that I was myself engaged in. Accordingly, my presentation dealt with this situation in a sometimes brash and polemical, sometimes shrill and hapless manner. It goes over some of the same material that has been presented above, this time in an argumentative rather than descriptive manner. I offer it here as the document of one attempt to move beyond ethical reflection under conditions of incredibility:

160 years after Hiroshima and Nagasaki, 20 years after Chernobyl and Bhopal, and only a few years after BSE we are living in an age of incredible confidence in what technology can achieve. This confidence is shared by those with great expectations and those who are looking anxiously toward the future. Even skeptics regarding the benefits of emerging technologies are quick to believe – if only for the sake of argument – the many claims of impending technological breakthroughs.

The *Guardian*, for example, is not known to be an uncritical newspaper. Nevertheless, under the headline "There is no stop-button in the race of human re-engineering" it suggests until the very last lines that future technology comes upon us like an irresistible force of nature rather than a political and cultural construct. It also provides an assessment of the prediction that in 30 years and with a life expectancy of 110 years we will routinely use memory enhancements and brain implants. Curiously, however, the assessment is not based on consultations with cognitive scientists or medical researchers. The predictions are to be anything but far-fetched simply because they are being made:

> Sound far-fetched? It's anything but. This is the most conservative of a range of scenarios about the possibilities of ‘human enhancement’ that have prompted fierce debate in the US and are exercising many a scientist's mind around the world. (Bunting 2006)

Similarly, James Wilsdon of DEMOS can hardly be accused of being an uncritical technophile. But even as he champions the cause of a democratic and open debate of the subject, he is under the spell of the "if and then" upon which this entire conference is premised. Yes, he writes in the *Financial Times,* what we are talking about belongs as of now on the pages of science fiction but we are already beholden to the moment when science fiction moves into the laboratory and just as surely into the marketplace. In the space of a half-sentence, Wilsdon traverses with amazing speed the huge distance from science fiction to commercial fact, from now to a remote, perhaps hypothetical future: “Yet as the technologies for human enhancement start moving from the pages of science fiction into the laboratory, and eventually into the marketplace...“ (Wilsdon 2006).

Finally, many skeptics and enthusiasts alike are transfixed by a graph or logarithmic plot that shows the exponential growth of technology as it heads toward a so-called singularity. It is fair to say, I believe, that this graph has credibility neither among historians of technology nor among statisticians. If it nevertheless enjoys credibility and considerable popularity, this is largely due to its apparent similarity to Moore's Law that describes the rate of progress over the last few decades in the semi-conductor industry but appears to have transmuted into something like a law of nature that can be extrapolated into the future. The graph is credible and popular also because it comes recommended by an ingenious wizard who has invented the flatbed scanner. We would be less credulous, I believe, if all we knew about its author is that he seriously believes to have physically aged only 2 years over the course of altogether 16 years. Either way, it would help if he had more of a historian's sensibility or a critical awareness of statistical inference.

I will argue that the spell of the "if and then" needs to be broken, that we must not be transfixed by visions of a future that is quickly descending upon us "with no stop button." Against this kind of bewitchment of our mind we need a reality check – and we need this not because there is something wrong with thinking about a remote technical future or with using science fiction scenarios to question who we are or wish to become. Indeed, there is nothing wrong with that in the context of philosophy and human self-reflection, and hardly a more urgent question to ask than: "Suppose you were free to choose your body and mind, would you choose yourself more or less as you are?"

If we need a reality check, it is also not because the new and emerging technologies aren't powerful enough to transform the world as we know it and implicate us in this change. They are, and I will give examples of this. We need the reality check in order to take responsibility for technological change, in order to see where the action is and where political action is possible, even required. We thus need the reality check because we can't afford to be barking up the wrong tree and be distracted from the issues before us.

I will argue for this reality check without claiming that I can provide it all by myself and from the vantage of a philosopher of science. At best, I can share some of my observations of scientific and technical developments, the observations of an outsider who attends numerous conferences on nanotechnology and converging technologies, and who served as rapporteur of the European Commission's expert group on converging technologies, authoring the so-called C*TEKS*-report that provides an alternative vision to the US-American NBIC Convergence for Improving Human Performance. In this respect – and in this respect alone – I am here speaking as a kind of European counterpart to William Bainbridge. I will argue that the C*TEKS* report frames the enormous potential of enabling technologies more appropriately – and that this can be seen especially in what is says and doesn't say about "Tomorrow's People."

I recently attended two scientific meetings devoted to the state of the art regarding brain-machine interfaces and found out that only 4 years ago it was possible to transmit by way of mental exertion no more than 2 bits per minute and that one has since then already achieved the capability of typing a 5-letter word per minute. What are we to make of this?

Some, of course, are plotting these values on a curve that looks once again a lot like Moore's Law. Yes, if this rate of progress continues, we will soon achieve a communication speed similar to natural speech – and surely will then go on to exceed that speed. Others will point out that this technical development does not depend on the number or size of electrodes, their sensitivity to brain activity, or the size of data streams. It depends crucially on the ability of humans to control this technology through proper and efficient training, concentration, and mental effort. Very ordinary human beings must be taught to create brain impulses that can be translated into the desired cursor motions. This work is so strenuous and fatiguing that it is likely to benefit only those who have lost all muscular control. As long as you can move any muscle at all, using it is likely to be the more effective interface between the human body and mind and the outside world.

But whether or not the transmission rates can be greatly improved, this has nothing to do as of yet with controlling a device by way of thought. A thought, after all, is something that has content or meaning. For several hundred years, science has pursued the dream of localizing thoughts as physical objects in the brain. I am not sure that any progress has been made in this regard. Some of the more promising theories of language and thought suggest that there can only be shared meanings and that a thought is therefore a social thing that exists not in the brain of individuals but among the minds of many. I don't know how such debates will be settled, but I do know that the settlement is not in sight. And as long as this is not settled, I very much doubt that regarding mind-machine communication or the nature and physical basis of thought we are on a technological trajectory like Moore's Law. There simply are no current trends that – if continued – entail the solution of these problems.

In light of all this it is strange, if not ridiculous that serious ethicists move as swiftly as James Wilsdon when they extrapolate from a very early stage of brain-machine-interface technology to the ethical concern that this technology might be used to read our minds and thus violate rights to privacy. This kind of ethical reflection is neither cutting-edge nor far-sighted, it has merely lost traction in a changing world in which a bewildering multitude of scenarios clamors for attention, in which the old hierarchies of intellectual authority and expertise have given way, in which we therefore find ourselves on shifting grounds. (I will spare you an analysis of how deeply such uncertainty and confusion reaches into the scientific communities themselves as multi-disciplinary research must seek funding by forging alliances that are based on trust.)

Now, if we shift grounds and look elsewhere – as we should – we find very powerful machine-brain-interfaces. Rather than try to get signal and perhaps thoughts out of the brain, deep-brain-stimulation produces effects in the brain and already exerts amazing and disturbing influence on human motor control but also on mood. It helps patients with tremors or debilitating depressions to resume on the flick of a switch a nearly normal life. It does so by performing very small lobotomies. There is nothing particularly sublime or marvelous about this. Instead of liberation and transcendence it invokes the idea of technical dependency and even the scenario of remote-controlled humans – of which we would hardly say that they are enhanced or that they possess extended powers of self-determination, even if we placed the remote-control in their own hands. Here, the emerging attention of ethicists is called for. We need their close engagement with particular cases as well as their discussion of general principles.

The proposed reality check does not stop here. Aside from considering limits of technical feasibility, we must examine unquestioned assumptions about human wants and about likely commercial developments, assumptions along the lines of "if this were available, wouldn't everybody want it?"

It is interesting to note that discussions of better humans are rarely clear about their subject matter. Are we talking about an improved human species because humanity can do better than produce an occasional Galileo, Einstein or Ray Kurzweil? The question is barely intelligible and impossible to answer. We don't have the vantage point from which to assess whether the human species needs a philosopher like Immanuel Kant who can paint and design like Leonardo da Vinci. Would we even notice this person? Would humanity be better off if more of us were like Mahatma Gandhi or Martin Luther King – or would a crowd of Gandhis and Kings start beating up on each other? While it is hard to envision a human species that is intellectually and morally better than itself, it is much easier to think of our personal mediocre selves in need of improvement: If only I could compose like Beethoven, play the piano like Glenn Gould! Of course, for intellectual, artistic, and moral improvement we have the old-fashioned enhancement strategies of education, of interest and immersion, of ambition and appropriation.

There is a third way of thinking of the subject matter of enhancement debates, namely to think neither of humanity as a whole or of our own mediocre selves but to assume a detached engineer's perspective on human minds and bodies as more or less well-designed technical products. Of course, when we then go on to ask how these already pretty highly developed and most complicated artefacts can be improved, we end up with the same stereotypes that dominate most commercial product development. Could it be more durable, stronger, lighter, faster, and of course smaller, cheaper to make, expensive to sell, and ever so easy to replace?

And thus the predominantly male community of human engineers or transhumanists begins focusing on features mainly of physical and mental prowess: live longer, jump further, see more, process extra information extra quickly, sleep less, extend your reach. We then go on to discuss these stereotypical engineering features in the equally stereotypical terms of individualist or consumer rights: If this becomes available who shall be able to prevent me from having it?

Again, I will recommend that we shift grounds. As for extending human reach, we are often presented with the spectacular findings of Miguel Nicolelis and Kevin Warwick. Nicolelis had a computer analyze the brain signals of a monkey and thus have the monkey move a robotic arm. Kevin Warwick experimented on himself, placed electrodes in his arm to move that of a robot. I hope you appreciate the humor or perhaps the rather crude conception of extending human reach in both of these news items – we are to be amazed that the monkey moved the robotic arm 950 kilometers, Kevin Warwick even several thousand kilometers away.

As before with the shift to deep-brain-stimulation and the prospect of remote controlled humans, a small shift of perspective proves sobering while giving rise to a host of more immediate, more relevant questions. When I see recent advances in robotics, in the development of data gloves and smart environments, when I am told that surgeries can be performed already at an arbitrary physical distance between patient and surgeon, why should I be impressed that Nicolelis and Warwick can electronically transmit invasively obtained data of correlations between brain or nerve signals and muscular action not only 100 feet or 2 miles but across the Atlantic Ocean? In light of the fact that data glove technologies will be further improved, refined, adapted to commercial applications long before I could even contemplate the acquisition of a somehow useful implant, the economic prospects of invasive enhancements appear doomed from the start.

Now, if you take a look at the previously mentioned C*TEKS* report to see what it says about physical enhancement technologies, you will find very similar considerations and very similar stories told. It raises the question whether someone with access to wrist-watch-sized voice-controlled expert systems and data banks would opt for a brain implant. It suggests that brain-machine interfaces can become a source of revenue and contribute to national prosperity in just those economies that are willing to invest in highly specialized and therapy-intensive treatments of rather small patient populations, and that they will fail where mass-marketable spin-off products for consumers are expected. And most generally, it recommends that science policy attends also to the limits of technical feasibility, suggesting for example that one should scientifically scrutinize the all too naive assumptions, if not (citing Dan Sarewitz) "conceptual cluenessness" about thought and cognition that underwrites the US-report on NBIC convergence. Along the same lines, a committee of historians and statisticians should produce a critical assessment of Ray Kurzweil's thesis about exponential growth. Also, as Jürgen Altmann has urged, we need an Academy report about the Drexlerian vision of nanotechnology – is molecular manufacturing a real possibility or not? Finally and most generally, we need scientists and engineers who have the courage to publically distinguish between what is physically possible and what is technically feasible. As a citizen, I am overtaxed if I am to believe and even to prepare for the fact that humans will soon engineer everything that does not contradict outright a few laws of nature.

By endorsing a process of agenda-setting for societally beneficial converging technologies research, the C*TEKS* report rejects the notion that we need technology to fully realize human potential. This notion posits an imagined ideal of what humans can, should, might be capable of, it then takes this ideal as a norm or teleological aim towards which technology must aspire. This notion therefore combines three undesirable features: 1) It defines the purpose of technology from the unavailable vantage point of some future humanity, 2) it proclaims an inescapable technological destiny that has no stop-button and leaves no space for political deliberation, and 3) it has the further immediate effect of instilling in Ray Kurzweil and his contemporaries what Günther Anders has dubbed the promethean shame of being born too early, even of being born and having to die at all, the shame of those who are presently alive and who view themselves as defective in light of abstract and idealized visions of a future humanity. Jean-Pierre Dupuy and Gregor Wolbring have drawn attention to this immediate effect of the notion that human potential is defined as something that we do not actually have and that only technology can realize – they will probably do so also at this conference.

By rejecting this notion, the C*TEKS* report replaces it with another notion: We need social innovation to bring out the potential of technology. This notion informs one of the report's conclusions, namely that instead of engineering *of* the body or *of* the mind, converging technologies research should be dedicated to engineering *for* the body and *for* the mind. This maxim owes not just to ethical conservativism or to considerations of technical feasibility. It also reflects that engineering *of* body and mind represents inefficient and unoriginal uses of technology. Such an approach is limited to work on one individual, one customer, patient, or consumer at a time. This individualism shortchanges us of what we can achieve through changes to the infrastructure or environment that might enhance human decision making and improve human interaction on a societal level – and which also happens to be quite within near- and medium-term technical reach.

The C*TEKS* report thus effects another shift of attention from individually focused human enhancement technologies to smart environments, ambient intelligence, ubiquitous computing – to cite just some of the labels for this family of applications. However, you will not find in the report a statement to the effect that these technologies are good while the envisioned physical enhancement of individuals is bad. On the contrary – because the idea of smart environments better utilizes the transformative potential of technology, because it will prove the more powerful, deeply pervasive technology, because it will structure our actions and therefore shape our thought-processes, our scope of freedom and responsibility, it demands our full attention. To the extent that we are the products of our products, that the world we make defines also who we are and what we can do in this world, we have to pay attention especially to questions of empowerment and deskilling, to new ethical questions that emerge as they did in the areas of human reproduction or death and dying, and to the ways in which our value-systems must and will adapt to this new world of our making.

We can break the spell of the "if and then" by focusing on the far-reaching issues right before us, by critiquing currently funded programs and unimaginative visions like that of "better humans," by adapting technological capacities to societal needs, by learning all we can about RFID-chips and their intended applications, proposing legislation on how to constrain those applications. If we do so – and current consumer activism on RFID chips testifies to this – we may find that we can do more than prepare ourselves for the inevitable or to discuss a technological development without stop-button. Instead, we may be able to shape the emerging smart environments just as users are shaping quite considerably the virtual environment better known as the World Wide Web.

As you can tell, I am emphasizing the fundamental distinction between socially embedded or infrastructural technologies like public transportation or energy generation, storage, and distribution, and individualized consumer technologies like Viagra and other performance enhancers or happiness-makers that require as their prerequisite a competitive world which produces intense pressures and unhappiness in individuals. By insisting on this distinction I am disagreeing with arguments that try to lump all technology under the label of "human enhancement." Arthur Caplan maintains, for example, that all technology improves and transcends our nature:

> That’s what agriculture is. That’s what plumbing is. That’s what clothes are. That’s what transportation systems are. They are all attempts by us to transcend our nature. Do they make us less human?

Caplan's off-handed remark is plainly wrong on three counts. First, agriculture, plumbing, and transportation systems did not seek to overcome or transcend our nature, that is our physical and cognitive limits, but to change nature or the world such that it would be more manageable for creatures with our physical and cognitive limits. Secondly, to the extent that agriculture, plumbing and transportation systems shape the world we live in and to the extent that we are the products of our products and of the world we make, our human values, perhaps natures are implicated in these technologies. Yes, our values and identities are changing alongside our technologies. But surely not in the direction of transcendence. As we acquire control by assimilating the world out there to the physical and cognitive requirements of our minds and bodies, we do not just liberate ourselves and extend power but also create new dependencies, new kinds of ignorance, new problems even of human or ecological survival. Finally, Caplan's remark glosses over the basic and intended novelty of current visions of human enhancement, molecular manufacturing, global abundance: As we rush towards and finally reach the "singularity," we are to be liberated, after all, from such tedious and stupid things like agriculture – as we extend our capacities and perhaps find a way to differently metabolize the air we breathe, we will no longer need ingenuity for achieving quite a lot with our rather modest cognitive and physical means. Let us not conflate, therefore, a transhumanist interest in technology for individualized human enhancement, with the tradition of enhancing ourselves through education and ingenuity and thus to use technology in order to better adapt the world to our interests and needs. And in light of James Wilsdon's statement that the transhumanist impulse might be progressive, let us not be too generous regarding our expectation of social progress emanating from a pursuit of happiness that basically consists in more experience, more fun, more self.

This session of our conference is devoted to "World Regional Perspectives" and it appears that, indeed, I am articulating a European as opposed to US-American approach: Instead of individualistic consumerism I recommend technologies with a defined societal benefit. In the place of technology as an avenue for human salvation and the achievement of transcendence, I view technology as a highly negotiated and ultimately political construct. Committed neither to ethical conservativism nor to economic neo-liberalism, I am recommending a democratic process that integrates social and technical change. And instead of focusing on so-called second and third generations opportunities and risks that lie in a distant future, I follow the example of the UK's Royal Society and Royal Academy of Engineering report on nanotechnology by foregrounding the more immediate prospects and challenges.

And yet, I fail to see in all of this a specifically European regionalism. The science-based development of technology is rooted in a tradition of truth-seeking, criticism, and enlightenment, the same tradition that produced the political constitutions not only of the US and Europe. The critical recognition of limits of feasibility, the creation of spaces for political deliberation and agenda-setting, the expectation of societal benefit from public investment hardly reflect regional commitments but are valued in all parts of the world. The question is only whether these commitments prevail in the formulation of sound science policy or at this World Forum on technological challenges.

[[Zusätze/Fußnoten? – ein zusätzliches Problem bei der Frage nach dem Menschen, der wir werden wollen??

This is not the place to critique the flawed methodology behind this curve. Here, it is interesting to note that it transfers to the whole of technological development Moore's Law which for the last few decades has functioned as a roadmap in the semi-conductor industry and which has everyone guessing how much longer it can be upheld. Also, we might just ask ourselves whether a person who was born in 1880 (before the telephone, automobile, and electric lighting) and died in 1960 (with space travel and nuclear power) really witnessed slower technological development than one who was born in 1920 and died in 2000 – and got to see lasers and semi-conductors and an ever-faster turn-over of electronic updates?

Consider, in contrast to all of this, the case of global warming. Here, we have been careful to convince ourselves that the trend is real or at least highly plausible – there are no assumptions about entirely new capabilities emerging or about sudden leaps in basic scientific knowledge, there is no denial of technical or scientific limits, there is no blanket appeal to the unfathomable deliverances of nano- and biotechnologies and their convergence. Also, after we persuaded ourselves of the real danger of global warming, we started looking for cultural and political interventions along with highly specific technical programs that might slow down, perhaps stop the trend. We do not sit down to develop an ethic or economic forecast for a futuristic world-under-water.

Fixated upon unlikely future scenarios of technologically enhanced individuals, we may actually blind ourselves to the transformative potential of current technical developments. Global warming is one of them, the creation of smart environments through ubiquitous computing technologies is another. Against the prospect of life in a greenhouse with memory and intelligence stored in the environment, the notion of an individual human being with a memory-enhancing brain implant appears not only less likely but also pathetically irrelevant and touchingly old-fashioned.]

References.

Amato, I. (1999): *Nanotechnology - Shaping the World Atom by Atom,* Washington, National Science and Technology Council, Interagency Working Group on Nanoscience, Engineering and Technology.

Nick Bostrom (2006) "Welcome to a World of Exponential Change", in: Paul Miller and James Wilsdon (eds.) *Better Humans? The politics of human enhancement and life extension,* London: DEMOS, 2006, pp. 40-50.

Arthur Caplan (2006) "Is it wrong to try to improve human nature?", in: Paul Miller and James Wilsdon (eds.) *Better Humans? The politics of human enhancement and life extension,* London: DEMOS, 2006, pp. 31-39.

Sendung im Bayerischen Rundfunk (2003): „Das Nanoschnitzel Vision und Wirklichkeit in der Nanotechnologie“, in der Reihe *Faszination Wissen,* Buch und Regie: Christian Friedl, gesendet 23.10.2003, 19:30 Uhr

Binnig, Gerd (2003) „Aus dem Nichts“, München: Piper 2003

Bruland, Kristine und Mowery, David C. (2004) „Innovation through time“ in: Fagerberg, Jan, Mowery, David C., Nelson, Richard R. (eds.) *The Oxford Handbook of Innovation,* Oxford: Oxford University Press, 2004

Madeleine Bunting (2006) “There is no stop button in the race for human re-engineering” in: *The Guardian*, Monday January 30, 2006

Crandall, B. C. (1996) “Nanotechnology – Molecular Speculations on Global Abundance”, London: MIT Press, 1996

Drexler and Smalley

European Commission (2003): *FP 6 "Integrating and strengthening the European Research Area” – Thematic Area 3: Nanotechnology and nanosciences, knowledge-based multifunctional materials, new production processes and devices – Work Programme,* edition December 2003.

European Commmission (2004): *Communication from the Commission: Towards a European Strategy for Nanotechnology*. Luxembourg, S. 10.

Hanson, Robin (2006): title of lecture?? at the Tomorrow’s People conference in Oxford, March 2006.

Harris, John (2006) like Hanson

Jones, Richard (2004) “Soft Machines”, Oxford, Oxford University Press, 2004

Kurzweil, Ray (2000) “The Age of Spiritual Machines”, New York, Penguin Books, 2000

Kurzweil, Ray (2005) “The Singularity is near”, New York, Penguin Books, 2005

Miller, Paul and James Wilsdon (2006) "Stronger, longer, smarter, faster" in Paul Miller and James Wilsdon (eds.) *Better Humans? The politics of human enhancement and life extension,* London: DEMOS, 2006, pp. 13-27.

Nordmann, Alfred (2005) *Noumenal Technology*

Nordmann, Alfred (2006) *Shifting Grounds: From Tomorrow's People to Today's Questions* Lecture at the Tomorrow's People (?? – like Hanson) conference in Oxford, March 2006.

Phoenix, Chris (2006) "Inapplicable Intuitions," *C-R Newsletter* no. 43 (July 30, 2006), Center for Responsible Nanotechnology.

Mihail C. Roco and William Sims Bainbridge, eds. (2002) Converging Technologies for Improving Human Performance, Arlington: National Science Foundation/Department of Commerce.

Sarewitz, D. (2005) "Will enhancement make us better?," *Los Angeles Times*, 9 Aug 2005.

Stoermer, Heinz (2005oder 2006) Laptops auf Bäumen (oder so) Süddeutsche Zeitung 9. (Dez 2005??)

Weber, Jutta ??

James Wilsdon, *Financial Times,* February 7, 2006

Workshop (2006): *State-of-the-Art and Socio-economic Aspects of Brain-machine (BMI) & Brain-computer Interfaces (BCI)*. Unpublished draft minutes.