Future, Humans and Technology

Michael Freimüller

This essay is on technology ethics, more specifically on the question, how we should deal with technological advances. It incorporates two very different views, on the one hand the view by Bill Joy, as presented in his essay "Why the Future Doesn't Need Us" (Joy 2000), which I call the pessimistic view. On the other hand we find John G. Messerly, who, in his essay "I'm glad the future doesn't need us: a critique of Joy's pessimistic futurism" (Messerly 2003), represents a, what I shall call, optimistic view. It is my goal to show, that neither view is entirely correct and that instead of going to extremes, we would do better to listen to both sides and find a middle ground that serves our interests better.

I will begin with a very concise summary of both points of view. In his essay "Why the future doesn't need us", Bill Joy argues, that technological advances, if left unchecked, may come with "unintended consequences". He bases his argument on a thought experiment, where "computer scientists succeed in developing intelligent machines that can do all things better than human beings can do them". I will use the term "AI" (artificial intelligence) as substitute for "intelligent machines". Given that thought experiment, Joy argues, that "the fate of the human race would be at the mercy of the machines", because "the human race might easily permit itself to drift into a position of such dependence on the machines that it would have no practical choice but to accept all of the machines' decisions". We can see this phenomena play out in small scale with modern technology. With online shopping, we permitted ourselves to fall into a position of dependence, because it is more convenient to go online and order precisely what we want, instead of driving to a shop and trying to find what we were looking for. I think, that we can call this a "light dependence" on technology, because in the absence of that technology, our lives (or at least most of them) wouldn't be in danger. Stronger dependency on technology is found in computer technology as such. Imagine a scenario where, due to a solar flare, all electronics went offline for a couple of weeks. In this scenario more than just a little inconvenience is to be expected. Because much in our lives depends on computer technology, our lives would be (severely) impacted. We would still be able to survive, because we could revert to "the old ways", but much of our daily comfort would be gone. But if, as Joy postulates, we fall into a dependence on AI, where AI plays a significant role in our lives and takes over many tasks in our lives, our infrastructure etc., it becomes clear, that our dependence on AI would result in catastrophic outcomes, given a blackout.

Of course, this is only one sort of dependency, and Joy acknowledges that this is only one possible outcome. The other form of dependency is the dystopian view, that, because "human work (...) will be superfluous", the "masses" will be at the mercy of the "elite", which, in the end, will engineer life in a way that it "will most certainly not be free", instead human beings will be "reduced to the status of domestic animals". In this, undebatably pessimistic, scenario, AI would not be in complete control, but will be a tool, used by the "elite", to control the "masses" and thus limiting freedom excessively. I think, that this scenario, just like the one discussed above, is to be taken seriously. To dismiss the possibility of this happening would be reckless, because it would ignore the history of humanity, which seldom shows great respect of the ruling class towards the ruled masses. The bottom line, I think, is that a greater discussion is necessary to decide about the ownership of technology of as great a scope as AI. To envision a future that incorporates AI in a scope as predicted by Joy, would require us to not only develop the technological means, but the social ones too. While here is not the place to defend this theses of mine, I think, that the connection between technology and human beings is much closer than we think it is. It goes as far as seeing humans not as "homo sapiens", but rather as "homo technologicus", for a lack of a better term. If we look at the rapid changes in technology in the recent years and the changes in human behavior, that accompanied those changes, we can view humans as beings that change themselves by changing their technology. Think about the way modern communication technology changed our lives and the ways we interact with each other, to a point, where, e.g. in modern dating, technology becomes the primary means of communication, instead of one-on-one talking. If that is the case, we can make a point in thinking carefully about the technology we introduce in our lives and, furthermore, we can argue that, if we introduce a technology, we should think about the consequences it has on us and ask the follow-up question, how we should change, to make the technology best serve us, and therefore produce the best possible outcome for our new understanding of ourselves.

Joy concludes his discussion by arguing, that, in the thought experiment of his, all possibilities "are (...) thus either undesirable or unachievable or both". Therefore, he proposes to "limit development of the technologies that are too dangerous, by limiting our pursuit of certain kinds of knowledge". As already hinted by the discussion above, I don't agree with his conclusion. While I agree that there are dangers on the road ahead, burying our heads in the sand won't deflect them. First off, it seems impossible to me to completely stop either the development or the pursuit of knowledge of AI technology. While some people might adhere to this call, others, I will call them rogue scientists, will continue their studies. If those rogue scientists succeed in their undertaking, they will have all the knowledge, while the rest of the world will be oblivious to the workings of AI, and will thus be depending on those rogue scientists. Therefore it would be better would the development and study be open to the public, because technology such as AI, which has a large impact on our lives, should not be developed and studied behind closed doors, but in the open, to

allow for independent control, and understanding of the technology.

We can summarize the pessimistic view on technology as the view, that technology can pose a significant threat to humanity and thus should be either strictly controlled in development, or banned completely.

I will now turn to the opposite philosophy. John G. Messerly, in direct response to Joy, in his essay "I'm glad the future doesn't need us: a critique of Joy's pessimistic futurism", argues, that technology might improve our lives, maybe even in a dramatic way, and we should therefore not worry about unintended consequences. Same as with Joy, Messerly's core argument concerns the unintended consequences that might result from technology. He argues, that "it is hard to quibble about the existence of unintended consequences", because if "the future is unknown, some consequences are unknown". Because we cannot know about the unintended consequences of any technology, Messerly concludes, that Joy's conclusion to "cease and desist in the research, development, and use of 21st-century technologies" cannot be supported, because he cannot know, that these unintended consequences will occur. Instead, he argues, that "it might well be that newer technologies will lead to a safer world".

In what follows, Messerly discusses all of Joy's arguments and refutes them by implicitly relying on his base premise, that the future is unknown, therefore it is not sure, what will happen. As technically true that premise is, as unsatisfying it appears to me. Yes, it is true, that we don't know, what will happen. Yes, it is true, that "we may increase our survival chances by switching control to more failsafe robots designed and programmed by our minds". Yes, sometimes we "*underestimate* our design abilities". But all these arguments look at the issue from the opposite side. What is unsatisfying to me, is the ignorance on potential threats that might emerge. While I don't fully agree with Joy on all his arguments, I also don't think, that we should dismiss them so easily. After all, there is wisdom to be gained by picturing the worst case scenario, even if it never occurs. To assume, that we don't know, therefore we shouldn't assume the worst, but instead assume, that things will turn out okay, or even better, falls into the same fallacy Messerly pointed out. We don't know, therefore we don't know if it will be good or bad. Only because we might have a more optimistic point of view, doesn't mean, that pessimistic scenarios are to be dismissed. The only real conclusion that Messerly can draw from his base premise, that we cannot foresee the future and its consequences, is that we should withhold any opinion on the subject, because if we cannot meaningfully speak about pessimistic outcomes, we cannot meaningfully speak about optimistic ones.

As optimistic as Messerly is, and as easy as he dismisses Joy's conclusion, that we should stop or limit the development and study of certain technologies, it seems as if Messerly wouldn't always agree with himself. In his essay "How Computer Games Affect CS (and Other) Students' School Performance" (Messerly 2004), Messerly argues, that video games severely impact students' lives due to their addictive nature. He argues, that gaming experience leads to the collapse of one's social life, gaming seldom offers positive lessons and games don't round

out a contemporary computer science education (Messerly (2004) p. 31). He concludes with the plea, that "we can only hope (that) gamers begin to recognize that the real world holds much more reward (...), promising more positive experience, knowledge, joy, and love" (Messerly (2004) p. 31). I think, that we can reconstruct Messerly's argument to fit into the general discussion of this essay.

Video games, I am sure you will agree, are technology. They require computer technology to work, because they are programs and assets that are only possible with computer technology. The goal of video games is to entertain people. That some people develop addictive dependencies on video games, is (in most cases) not intended, therefore these addictions are unintended consequences. Therefore, Messerly argues, that it would be better, would students not play video games, because of their unintended consequence of addiction, and instead rely on other means, that is, the "real world". Now we might ask two thins: First, if video games showcase, that technology can have unintended, negative consequences, why should we look at technology in an all positive light? Second, considering the term "real world", where does it begin and end?

The first question, if video games showcase, that technology can have unintended, negative consequences, why should we look at technology in an all positive light, seems to me the weak point of Messerly's argument. It seems to me, that Messerly distinguishes between *types of technology*. Maybe Messerly assumes, that there seems to be a category of technology where unintended consequences are more unlikely, or maybe even preferred, and another category, where those unintended consequences are making our lives worse. The issue is, as Messerly says, that we don't know, what the unintended consequences of technology are. Therefore we would have to look at only the intended consequences. Here, we could argue, that video games clearly are examples of bad, unwanted technology, because they served no purpose for humans. They "merely" entertain players and remove them from the "real world". Whereas, a chip that is implanted in the brain, that improves cognitive function, is "good" technology, because it enhances humans in a productive way. That the chip can go bad and cause damage to the brain, or turn into a controlling chip, must be irrelevant on that account, because we cannot foresee this (although we can clearly imagine this outcome). Analogous, the same is true for video games. When we design video games, we cannot foresee, that some players will become addicted to them. Therefore, we cannot prevent the development of video games on that premise, but rather because the core idea of video games, the intended consequences, seem to be contrary to what we want. But who is to decide, what intended consequences we want to accept? Even today, where technology is not as enhanced as Joy sketches it to be in his thought experiment, we see, that there is next to no (public) discussion on the intended consequences of technology. The question: "Do we want to invite this technology into our lives, together with the intended changes it has on our lives?", is seldom, if ever, discussed. We would need an instance, a group of people, that decided what technology to allow and what technology to forbid. We can imagine a scenario where video games, with their non-productive use,

would be forbidden, while brain implants would be allowed. But this approach would be dangerous and reckless. Dangerous, because who is to say, that the decision that these experts make, are in the interest of the people? And reckless, because ignoring unintended consequences is like a head-first jump into potential shallow water. It is not wise to do. It's like planning an undertaking and ignoring all potential causes of failure. Sure, they may be unlikely, but should we not talk about them? And if we see, that the chances of success are more unlikely than the chances of failure, are we wiser in taking the risk, or taking a step back and rethinking our choices?

The second question, concerning the term "real world", is a tricky one. Messerly argues, that Joy's fear of nanotechnology and the human enhancement project in general, that is, for example, the uploading of the human mind into a computer, or the merge of humans and computers, are unfounded, because human beings are always evolving and in case of standstill, where we stopped our development, many people would prefer death to such a standstill. But if we intervened with "normal" evolution and changed our nature with technological means, wouldn't this be an intervention with the "real world"? What, then, is the "real world", after all? Is it that thing that we were born within, the untouched world? Or is not rather that the "real world" is constantly effected by us, therefore turning into a fluid concept? If video games, that is, a type of technology, seems to take people away from the real world, with all its hardships, joy etc., to a place where these experiences don't exist, why is it, that Messerly seems to openly invite a future, where we change ourselves (and therefore, the "real world") in a way where we also maybe don't experience certain things anymore, either because we used technology to disable certain experiences (like suffering and pain), or because we became part of a giant supercomputer, where, for example, hardship is removed from us? How can Messerly argue, that it is bad, that video games remove one from the "real world", while arguing, that in other cases it is not a bad thing at all? On what basis can we decide, with which technology these removals from the "real world" are good, and which are bad ones? And in how far are we willing to stretch our definition of the "real world", to accommodate our technological changes? Video games are, in a way, the ultimate experience of a virtual reality. Any thinkable limitations that are put on us in the "real world" can be lifted by games. In a way, game designers are the gods of their creation. But here, again, Messerly would argue in two ways. He chides video games for their virtual reality that removes players from the "real world", while arguing, that he knows "no reason—*short of childish pleas not to play God*—to impede our increasing abilities to perfect our bodies, eliminate disease, and prevent deformity" (emphasis mine). The human enhancement project, taken to its most extreme form, is similar to the ultimate video games. Humans are being upgraded in a way that presents them with next to no limitations. But we can ask: Would this still be the "real world"? Would we not loose the meaning of the concept "real world"? And if Messerly argued, that in this case it would be a good thing, we can ask, what the difference between "living" in a video game and living as enhanced human being would be? If we removed all the

hardship from life? If we were able to remove any unwanted emotions? Maybe even reverse aging or trick death? If this were the case, how can we draw the line between technology that changes us for worse and better? How, if not by limiting at least the development of certain technologies, *until we have a better understanding of them*? My point is not to go the pessimistic route and stay at the same place, technology-wise. Rather I propose a view, where we change our view on technology and development as such. Instead of plunging head first into new technology, we should assert the changes it makes on humans as such. Maybe we should change our view on technology to be similar to medicine. Just as medicine helps the body fight of disease, so technology helps humans in various aspects of their lives. While there seems (or should be) no controversy over the fact, that the uncontrolled use of new, insufficient studied, medicine is reckless, so should be the view, that the uncontrolled use and distribution of unstudied technology is reckless.

Although I spent much of this essay discussing Messerly's optimistic point of view, I now want to take a step back and discuss both viewpoints from a neutral point of view. It is true, that the uncontrolled development of technology might lead to disastrous outcomes. We can - and should - imagine a scenario where technology turns rogue and takes control over our lives. We should imagine it, so that we can implement safety measure and introduce a responsible interaction with technology. But we shouldn't refrain from developing and studying technology altogether. Of course that implies, that we should neither be too optimistic nor too pessimistic. We should seek for the middle ground, where we understand ourselves as species that creates itself through technology, and not the other way around. We should start a discussion about the changes we would like to make and about limits that we wouldn't want to cross. This is one part of the ethical discussion of technology.

One point that we should always keep in mind, is the goal of technological advance. Why do we develop and study new technology? Is it an end in itself, or do we strive for a higher goal? One possible, and undeniably important, goal of technological advance seems to me the continuing survival of humanity. Given the fragile state we are in, considering dangers on our planet (for example: climate change and depletion of resources), and dangers from out of space (for example: asteroids, comets and solar flares), we should seek to find ways to make our survival possible for as long as possible. Technology is our only and best bet. Of course, there are dangers on the road ahead. A society that solely relies on technology can find itself in great peril, should the technology stop working. If all our knowledge is concentrated in technology and more and more of our lives are automated and taken over by AI, one blackout is enough to cripple all of life and endanger the survival of humanity. While there are great advantages in advancing technology, we shouldn't forget the old ways, that in case of catastrophes enable us to continue our lives, even if with less comfort than today.

References

Joy, Bill. 2000. "Why the Future Doesn't Need Us." https://www.wired.com/2000/04/joy-2/.

Messerly, John G. 2003. "I'm Glad the Future Doesn't Need Us: A Critique of Joy's Pessimistic Futurism." *ACM SIGCAS Computers and Society* 33 (2).

———. 2004. "How Computer Games Affect CS (and Other) Students' School Performance." *Communications of the ACM* 47 (3).