February 2022

Our Robots or Theirs?
Contribution to GTI Forum Technology and the Future

Deborah G. Johnson

Thinking about technology and the future is like thinking about an ocean liner moving in the sea and wanting to know its current destination or, more importantly, wanting to change its current course. The ship (the world we live in) is underway; no one can know with complete certainty where it will arrive, at least not in a long-term timeframe. That depends on so many and complicated factors, including the weather and conditions in the sea (nature, the earth), as well as decisions made by the captain, orders given to the captain by the ship owners based on their motivations and calculations, the mechanics of the ship, and the behavior of the crew, just to name a few. Although pretty good predictions can be made about the destinations the ship will reach in the short term, predictions farther into the future are fraught. Changing the direction of the ship is a daunting challenge requiring coordination and alignment of the many, many forces that keep the ship moving on its current course.

Many believe that technology is the major force directing us to the future. As many would have it, so-called "emerging" technologies will get us to a better place. For example, Bill Gates seems to believe, and invests in his belief, that new technologies will avert climate disaster.[1]

The problem is that technology is not an exogenous force: it does not follow its own course of development independent of other forces. Technology—technological artifacts and systems—is inextricably intertwined with the human social world in which it is embedded. The technologies we have now, and those that are on the horizon, did not come out of nowhere; they are products of social, economic, political, and historical forces, forces that continue to keep those technologies in place. Technology is not, at least not in any simple way, the steering wheel of the ship. Yes, new technologies have the potential to tweak the direction of the ship, but they

also have the potential and the momentum to resist change and keep things the same. That is because technology isn't just artifacts and material entities. The ship has material parts, but it is also people believing and behaving in a wide variety of ways. Think here of our energy systems and the institutions and cultural beliefs that keep us dependent on fossil fuels.

Technologies are sociotechnical; they are assemblages of artifacts together with human behavior, social institutions, social practices, and social meanings. Let's take one of the seemingly more fanciful predictions about technology and the future: intelligent humanoid robots. Technologists are hard at work developing robots that look and act like humans. R&D efforts seek to develop materials that are humanlike in appearance, sound, and touch; AI is being developed to make such robots highly socially interactive. The ship seems to be moving to a destination point in which it may be difficult for humans to tell when they are interacting with another human being and when a machine. Indeed, some predict that humanoid robots with general intelligence and autonomy will in the future become so human-like that we will have to grant them some form of moral status or legal rights.[2]

Is a world in which we can't, and perhaps don't even want to, distinguish robots from humans inevitable? Will it be a better place? The inevitability question is wholly misleading. To get from here to there—as with a ship—is not a matter of sitting back and just letting "nature takes its course." A good deal of social and technological effort is necessary to keep the ship going; R&D efforts are what is bringing us humanoid social robots. Instead of thinking about the inevitability of such robots, we should be asking why the ship is headed in that direction. The answer is, of course, that there are a wide range of economic and market forces interested in this technology as well as an abundance of engineers and computer scientists who like this kind of challenge and hope to build their careers on achievements in this area. The point is that technology does not come out of nowhere. It is a human creation, a human product with meaning for humans. Even if humanoid robots continue to develop, what they will mean to us, how we will use them, whether we will recognize them as having moral status or not are all matters to be decided by humans. The big question is who is deciding and what are their interests.

To be sure, our relationship with computer technology is especially complicated because it is such a malleable technology. The malleability of computers and information technology has

led to its infiltration into nearly every domain of human life. Robots and AI are just one of the latest manifestations of human choices made about how to deploy the capabilities of computer technology. Some claim that current AI constitutes a special or even unique step in the expanding development of computer technology because AI constitutes systems that operate autonomously and learn as they go. The learning capacity means both that some AIs can make decisions that humans could not make (think here of AI systems detecting tumors that are not visible to the human eye) and that humans cannot always understand how an AI arrives at its decision (think here of an AI that detects fraud in millions of credit card transactions). While some argue that humans cannot be responsible for AI algorithmic decisions because humans don't and can't know how some AIs make decisions, the fact that humans cannot fully understand how AI algorithms operate means that mechanisms should be put in place to ensure that AI systems are not out of human control. Among other things, this means holding humans accountable for the behavior of AI systems.

Returning to the ship, the task of changing its course is daunting. Technology will play a role in where the ship goes, but if technology is developed in the same institutions that have produced it in the past, we have to be skeptical about the extent to which it can redirect the course of the ship. We have a chicken-egg problem here. We are not likely to change the course of the ship by the introduction of some disruptive technology unless the human institutions and motivations with which that technology is intertwined also change.

Endnotes

1. Bill Gates, *How to Avoid a Climate Disaster: The Solutions We Have and the Breakthroughs We Need* (New York: Penguin Random House, 2021).

2. David Gunkel, *Robot Rights* (Cambridge, MA: MIT Press, 2018).

About the Author

Deborah G. Johnson is Anne Shirley Carter Olsson Professor of Applied Ethics Emeritus in the Science, Technology and Society Program at the University of Virginia School of Engeineering. She has taught engineering students at Rensselaer Polytechnic Institute, the Georgia Institute of Technology, and the University of Virginia. Her research focuses on ethical questions related to computer/information technology and engineering ethics, STS theory, and technology policy. She published one of the first textbooks on computer ethics, the 1985 *Computer Ethics*, and co-edited the reader *Technology & Society: Building Our Sociotechnical Future*. She holds a PhD in philosophy from the University of Kansas.

About the Publication

About the Great Transition Initiative

The Great Transition Initiative is an international collaboration for charting pathways to a planetary civilization rooted in solidarity, sustainability, and human well-being.

As an initiative for collectively understanding and shaping the global future, GTI welcomes diverse ideas. Thus, the opinions expressed in our publications do not necessarily reflect the views of GTI or the Tellus Institute.