

## Interview with Philip Brey

**Jon Rueda Etxebarria**

La Caixa INPhNIT Fellow  
University of Granada  
[ruetxe@gmail.com](mailto:ruetxe@gmail.com)

**Txetxu Ausín**

Grupo de Ética Aplicada.  
Instituto de Filosofía, CSIC  
[txetxu.ausin@cchs.csic.es](mailto:txetxu.ausin@cchs.csic.es)

ISSN 1989-7022

Philip Brey is Professor of Philosophy of Technology at the Department of Philosophy at the University of Twente (Netherlands). He has been a keynote speaker of the *International Workshop on Controversies and Polarization on Disruptive Technologies*, that took place virtually and in Granada, on October 5th and 6th 2020. His key-talk was titled “Disruptive Technologies: Social, Moral and Ontological Consequences”.

Jon Rueda Etxebarria & Txetxu Ausín: “Interview with Philip Brey”, en Jon Rueda (ed.): *Tecnologías socialmente disruptivas* ILEMATA, Revista Internacional de Éticas Aplicadas, nº 34, 133-137

**1. The philosophy of technology has undergone many changes in recent decades. For instance, greater concern for practical issues and the empirical turn are outstanding examples in this respect. As a professor with extensive experience in the philosophy of technology, you are one of the protagonists who have witnessed this series of transformations. In particular, you are one of the most prominent authors in the ethics of technology. Although this field is in its infancy, this young discipline seems to have a bright future ahead of it. This is partly because society is increasingly demanding comprehensive analyses of the ethical and political aspects of technological developments. Is this a point of no return? Do you think that philosophers of technology have abandoned their ivory tower and are increasingly concerned with practical and societal issues?**

I think that philosophers of technology have always been concerned with practical and societal issues. However, until maybe the 1990s, the field has been largely theory-oriented. The reasons for this are, I think, first of all that the way that academic research was structured in most of the 20th century made it difficult to do applied research. But also, this was still a young field with many theoretical issues to sort out. If it immediately would have focused on practical issues, I think it would have lacked the theoretical background to do so. What I think has changed since the 1990s are the empirical turn, followed by what is sometimes called a policy turn. I think these turns were driven in part by the emergence of empirical studies of technology in the field of science and technology studies, and in part by the changing funding structure for philosophers. In-

creasingly, research funding became dependent on grants, and the grant programmes that emerged in the 1990s, at least in Europe, favoured applied research in philosophy of technology, especially ethics of technology, often with a link to policy as well. This is currently very visible in the funding scheme of the European Union, for example, where most of the funding in philosophy of technology is in applied ethics programmes under the umbrella of Responsible Research and Innovation. For at least the near future, the interest of society in funding applied ethics research in ethics of technology seems to persist. One important recent development is that many tech companies are now also hiring ethicists, for ethics policies and ethics research in their companies. This further exacerbates the shift towards practical and societal issues, although it also calls into question the independence of the field as more and more researchers are working for companies.

**2. Do you think that your own way of doing philosophy of technology has changed? What were your philosophical influences when you started your academic and research career? Are there any that have remained throughout your trajectory?**

In my own development, I started with theoretical philosophy (philosophy of mind, philosophy of language, epistemology) and then moved from philosophy of science to philosophy of technology, and ended up focusing most of my attention to ethics of technology. This was certainly supported by a lot of ethics teaching that I did early on in my career, which motivated me to address ethical issues in relation to technology. I think, also, that in my development, I moved from a more theoretical approach in the philosophy of technology to the empirical turn and then a policy turn. My empirical turn was aided by my interest in other disciplines and some training I received in science and technology studies (I studied with Bloor, Collins, Latour and Feenberg, amongst others, while at the University of California, San Diego), and my more recent policy turn has been stimulated by my participation in funding programmes of the European Union. I now find myself wanting to revisit the theoretical foundations of the field (philosophy and ethics of technology), and luckily, I was successful, along with a number of my colleagues, to secure funding for a very large grant to work on foundational issues for the next ten years.

**3. One of the best examples of society's demand for insights from the ethics of technology is the large number of research projects that are proliferating in this area. You have led various projects with European funding such as SHERPA and SIENNA, and you are also a partner of SHERPA. Recently, you have started leading a huge project about the Ethics of Socially Disruptive Technologies. This Gravitation Grant has a combined budget of €27 million. This is great news for the ethics of technology in general. What do you think you can contribute with this project to society? What is the goal of the project that motivates you the most personally?**

This grant, funded by the Netherlands Organisation for Scientific Research and by participating universities, aims to innovate the ethics and philosophy of technology over a ten-year period (2020-2027). I feel extremely lucky to be programme leader of such an ambitious programme. The reasons that I am so excited about it is that I felt that with the empirical turn

and policy turn in philosophy and ethics of technology, not enough efforts were remaining to give these fields solid theoretical and methodological foundations. Nowadays many people are doing case studies and other applied work, which does not always lead to innovative results. The field is at risk at being fragmented. You already see that with this new wave of ethicists of AI coming in, often people who have never worked in ethics of technology before, who bring conventional ethical approaches to AI, ignoring the lessons of decades of work in ethics of technology. In our programme, we aim to bring together some of the best people in the field, and grow a new generation of researchers who are able to do groundbreaking research and keep innovating the field. Part of the excitement is that we have ten years to develop the programme, and that we do not know where exactly we will end up, except that we will keep doing innovative work.

There are three a couple of themes in our research programme that excite me particularly, though. The first is the theme of conceptual disruption. We will be investigating how new technologies do not just disrupt and transform society, but often also challenge and transform our moral and ontological concepts, concepts like agency, human nature, liberty, equality, privacy and others. We will be studying how technological revolutions challenge these concepts, and what implications this has in turn for doing philosophy of technology, and philosophy at large. Another exciting theme is that of integrative methods for philosophy and engineering. While the philosophy and ethics of technology are becoming increasingly relevant to engineering, there are still few methods at the intersection of philosophy and engineering that integrate both. We will be working closely with engineers to further innovate in this area, taking up new challenges in design for values and ethics by design, but also to develop new ways of collaborating in other ways.

#### **4. What is a 'socially disruptive technology'? Can there be different types of disruption?**

Socially disruptive technologies are technologies that do not just change the specific domains or practices for which they were designed, but that change our life in a much broader sense. They are technologies that transform everyday life, social institutions, cultural practices, and the organisation of the economy, business, and work. They may even affect our fundamental beliefs, rights, and values. Historical examples of such technologies include the printing press, the steam engine, electric lighting, the computer, and the Internet. Now, we have AI, robotics, next-generation genomics, nanotechnology, and many others.

Indeed, new technologies can be socially disruptive in many ways, and part of our investigation is to investigate various dimensions of disruption. We are currently constructing a new impact assessment model for this, and we aim to investigate, particularly, disruptions to the basic functioning of society, to individual life and behaviour, to the natural environment, and to basic concepts and beliefs.

**5. In various publications, you have analysed the challenges of addressing the ethical issues of emerging technologies. Uncertainty about future technological developments is a major obstacle in this regard. The Anticipatory Technology Ethics (ATE) model is one of your contributions in this respect. Briefly, what are its main characteristics?**

The ATE model is intended to allow for comprehensive ethical assessments of emerging technologies. By a comprehensive assessment, I mean one that considers not only general ethical issues in a new technology field, but also various applied issues that are related to products and applications coming out of that field. Because we often do not know yet what these products and applications will be, we need to anticipate possible and plausible trajectories that the new technology can take. This is done using foresight analysis. ATE combines foresight analysis and ethical analysis to identify, analyze and evaluate ethical issues at three levels: the general level at which a technology is defined, the level of concrete artifacts and products, and the level of application and use, in particular contexts and by particular users. For example, if you were to assess the robotics, this approach would tell you what general ethical issues are in relation to it, but also the ethical issues in relation to specific types of robots, such as humanoid robots, social robots, swarm robots and micro-robots, and the ethical issues in relation to specific uses, such as robots in healthcare, robots in defense, and robots in law enforcement. But you can also choose to limit your analysis in ATE to only one type of artifact, or one domain of application—that is part of its strength.

**6. Many disruptive technologies may lead to major transformations in a future time that is difficult to foresee. Are you interested in combining the ATE approach with the analysis of socially disruptive technologies? What do you think are the technologies that will cause the greatest social impact in the medium term (say about 25 years)?**

ATE is well suited to be integrated in our new research programme, and I expect that to happen in the coming years. We are currently canvassing which new technologies we assess as being socially disruptive. We have identified a total of sixty so far. Some examples are: Internet-of-Things, Blockchain, sensor technology, augmented reality, genome editing, synthetic biology, brain-computer interfaces, carbon dioxide removal technologies, smart grids, and smart materials. Some of these, of course, will have broader scope and impact than others. Certainly, AI will have a big impact in the future, in combination with robotics and data analytics. Internet-of-Things will as well, as will genome editing, neurotechnology, new climate change technologies, additive manufacturing and new materials, to name some.

**7. This context of technological disruption is characterised by a growing interaction between humans and artifacts (human-machine-interaction), blurring the boundaries and enhancing our comprehension as socio-technical environments. What consequences does this growing hybridisation have, in your opinion, and how can it change our conception of moral agency and human identity? Linked to this, can we speak of moral progress? And what would be the most relevant elements of moral disruption in this new context?**

This is certainly one of the key questions for our new programme. While technology has always extended human agency, we are now coming to a point at which the relation between technology and humans gets so intimate, where technology is (semi)permanently attached on or inserted into the body that we are witnessing a hybridisation and cybernetization of humans that we have not witnessed before. This has potentially profound consequences for human identity, agency, and morality, and the very concept of human nature. We have to

ask for each new technology that has such an intimate relation to human beings: what are the consequences for these notions, and will they still allow us to speak of an independent, autonomous human subject capable of having its own thoughts and making its own choices rather than them being determined by technology? There is certainly a risk that this hybridization will lead to an abdication of moral responsibility by individuals and new responsibility gaps in society. This would not be moral progress. However, if technologies are designed to support human autonomy and liberty, then they may well end up supporting moral progress. These are very complicated moral and design decisions that would have to be taken in relation to new technologies. They especially play a role for neurotechnologies, AI, robotics, and wearable technologies, as well as gene editing technologies that may affect the way we think and act.

**8. Finally, there is no doubt that disruptive technologies have a huge transformative impact on politics as well. Are liberal democracies prepared to absorb this impact? How do disruptive technologies affect the political system? Do we need a new political philosophy for this new context?**

This is an increasingly pertinent question, and one that is central in one of the four research lines of our *Ethics of Socially Disruptive Technologies* programme. The disruptiveness of technologies can extend to the existing institutional and political order. And the resulting disruptions can be for good or for ill. So far, the Internet has had by far the greatest impact on liberal democracies compared to any other technology, as it has revolutionized the way that people collect and disseminate information and communicate with each other, processes that are vital to the functioning of any political system. The initial expectation was that the Internet would strengthen democracy by making information and communication more democratic. But the rise of commercial social media platforms in combination with data analytics and AI technology has come to support misinformation, radicalization and political manipulation in ways that now threaten liberal democracy. In our programme, we will examine these developments and ways to counteract them, as well as other ways in which emerging technologies disrupt and transform political systems.