

# Anticipatory Ethics for Emerging Technologies

Philip A. E. Brey

Received: 25 October 2011 / Accepted: 14 March 2012 / Published online: 4 April 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** In this essay, a new approach for the ethical study of emerging technology ethics will be presented, called *anticipatory technology ethics* (ATE). The ethics of emerging technology is the study of ethical issues at the R&D and introduction stage of technology development through anticipation of possible future devices, applications, and social consequences. I will argue that a major problem for its development is the problem of uncertainty, which can only be overcome through methodologically sound forecasting and futures studies. I will then consider three contemporary approaches to the ethics of emerging technologies that use forecasting: ethical technology assessment, the techno-ethical scenarios approach and the ETICA approach, and I considered their strengths and weaknesses. Based on this critical study, I then present my own approach: ATE. ATE is a conceptually and methodologically rich approach for the ethical analysis of emerging technologies that incorporates a large variety of ethical principles, issues, objects and levels of analysis, and research aims. It is ready to be applied to contemporary and future emerging technologies.

**Keywords** Anticipatory technology ethics · Emerging technologies · Uncertainty · Futures studies · Forecasting · Technology assessment

## Introduction

Different technologies find themselves at different stages of development and societal uptake. Some technologies have yielded many concrete devices and applications and are used by a many different people in a variety of contexts. For such technologies, ethical analysis has the benefit that many of the ethical issues have already been identified in society. For instance, a large variety of ethical issues in relation to the Internet have been identified not only by ethicists, but also by users and other stakeholders who run into them as they use or deliberate on the technology. Other technologies, however, are still emergent: they are at an early stage of development and have not yielded many applications and societal consequences. They are still largely, or fully, at the research and development (R&D) stage, meaning that they are still at the stage of research into basic techniques, or at an early stage of development which at most has resulted in lab prototypes and experimental applications but little or no serious products that are being used by ordinary users. These technologies will be called *emerging technologies*.

For technologies at the R&D stage, ethical issues relating to their use in society cannot be known reliably, as their impact on society lies in the uncertain future. At the research (R) stage, the stage of fundamental research, the focus is on basic techniques, principles and methods that can be used for later development of concrete devices or processes, whereas development focuses on the actual design and manufacture of devices and processes. At this stage, no knowledge may yet exist

---

P. A. E. Brey (✉)  
Department of Philosophy, School of Behavioral Sciences,  
University of Twente,  
P.O. Box 217, 7500 AE Enschede, The Netherlands  
e-mail: p.a.e.brey@utwente.nl

about possible devices or applications that may result from the research, so ethical reflection on future consequences may be wholly speculative at this stage. At the development (D) stage, the focus is on the design and manufacture of actual devices and processes. At this stage, more information is known about possible designs, but there is still considerable uncertainty about the devices and systems that will eventually gain societal acceptance, the ways in which these may ultimately be used, and the societal consequences that their use will bring. So at this stage, also, there is much uncertainty regarding ethical issues and ways in which these may be approached.

The question that is the focus of this essay is how we can identify and evaluate ethical issues for technologies that are still emerging because they are still at the R&D stage. With the accelerated pace of technological change in contemporary society, and the major impact that technology has on people's lives, early identification and evaluation of ethical issues is an important aim. Early identification can help users and other societal actors better prepare for future moral dilemmas, and can also help steer R&D or usage practices so as to avoid or minimize ethically undesirable consequences. Yet, so far very little research has been directed at developing sound approaches and methods for ethical analysis of emerging technologies. It is only in recent years that such research has seriously gotten underway. My aim in this essay is to review some of this recent work and to present a new, integrative approach for the ethical study of emerging technologies.

Ultimately, ethical assessment of emerging technologies concerns the question of what is good and bad about the devices and processes that these technologies may bring forth, and what is right and wrong about ways in which they may be used. Since at the R&D stage many devices, usage patterns and social consequences of the technology are not yet present, ethical assessment turns speculative, as it focuses on particular R&D activities and techniques and then projects possible devices and usage patterns which are then assessed ethically. Such assessments may then be used to make ethical recommendations for R&D practices themselves, so as to increase the likelihood that these practices yield morally desirable devices and uses. Or they may be used for policy.

The paper is structured as follows. In the next section, two approaches within the ethics of emerging

technology will be distinguished, based on how they deal with the problem of uncertainty about the future. In section 3, three recent ethical approaches to emerging technology will be discussed and critiqued, and it will be concluded that neither is fully satisfactory. In section 4, my own approach will be presented, which is called *anticipatory technology ethics* (ATE). I will present ATE as a promising new approach that builds on previous approaches, and I will provide examples throughout its discussion how it can be applied in practice.

### **Ethics, Uncertainty and Forecasting**

The central problem for an ethics of emerging technologies is that we do not know the future, and therefore do not know which ethical issues will play out once the technology is fully developed and entrenched in society. Because emerging technology is technology in the making, many questions about its nature, its future use and its social consequences are still undecided. For this reason, many ethical issues in relation to it cannot yet be identified or analyzed reliably. We can speculate about future applications and uses, but as history has shown, speculations about future technology are often way off the mark, meaning that we may end up exploring a misguided or irrelevant set of ethical issues.

The ethics of emerging technology therefore has to deal with an epistemological problem, the *problem of uncertainty* concerning future devices, applications, uses and social consequences [8]. The question is how it can deal with this problem in a responsible manner. On the one hand, it is to be avoided that ethicists lose themselves in idle speculation on future ethical issues in technology that in most cases turn out to rest on mistaken projections on how the technology will actually evolve. On the other hand, it is to be avoided that ethicists feel that they can say nothing about emerging technologies because they do not know which devices and uses will result from them. So the question is how ethicists can come to assessments of emerging technologies that are based on somewhat reliable knowledge of the future.

Two approaches are possible at this point, one more conservative and reliable, the other more uncertain and speculative. The first approach is to restrict oneself to ethical analysis of generic qualities of the new

technology that are likely to manifest themselves in all or most future applications of the technology and that are likely to present ethical challenges. For example, when nuclear energy technology was being developed it was known early on that however it were to be developed, there would be a problem of radioactive waste, which requires ethical deliberation. When genetic technology was being developed it was known from the beginning that it would involve the modification of genetic material, which was considered to be intrinsically morally controversial. So even when particular applications or uses are not yet known, it is often possible to identify generic ethical issues that are likely to manifest themselves as the technology progresses, and these can be discussed at an early stage. I will call this approach the *generic approach*.

A second approach is to speculate on future devices, uses and social consequences. This requires that ethicists either rely on existing forecasting studies or do such studies themselves. They can then use the forecasts to explore ethical issues. For example, ethicists can forecast that nanotechnology will yield applications for targeted drug delivery in the human body using nanoparticles, and that such applications will become widely available to both doctors and patients. They can then analyze ethical issues that are likely to occur when such devices are being used. I will call this the *forecasting approach* to the ethics of emerging technology.

The forecasting approach relies on predictive studies of future technological devices, uses and social consequences. Such studies are undertaken in two related fields. *Futures studies* is a field that aims to study what possible or probably futures may look like [1]. Futures research includes many different forecasting approaches, such as environmental scanning, causal layered analysis, the Delphi method and scenario methods. Some of these, like the Delphi method, rely on the consultation of experts in various fields, whereas others may rely on surveys, time series analysis, regression analysis, or simulations. Some work in futures studies focuses on technology forecasting. It forecasts future technologies, including the development spread of certain types artifacts, and optionally their utilization and social consequences that may result from their use. *Technology assessment* (TA) is a field that studies the effects of new technologies on industry, the environment and society, evaluates such effects and develops instruments to steer technology

development in more desired directions [5,12]. It makes such assessments on the basis of known or potential applications of the technology. Thus, TA in part relies on, and in part engages in, futures studies. Both futures studies and TA can hence be useful for forecasting the development of emerging technologies.

The forecasting approach has as an advantage over the generic approach that it is able to consider more ethical issues, by including not only those that are generic to the technology but also those that are specific to projected future devices and their uses. Its potential disadvantage is that its ethical assessments is based on forecasts that are to some degree speculative and that may be incorrect. However, to the extent that forecasts can be reliable, a forecasting approach will be able to anticipate many more ethical issues than a mere generic approach would, and would therefore be preferable. In the next two sections, therefore, I will focus on forecasting approaches. I will first look at three contemporary forecasting approaches to the ethics of emerging technology, which I will critically evaluate. In the section thereafter, I will then present my own approach.

## Critique of Existing Approaches

In recent years, forecasting approaches to technology ethics have been gaining attention, although few mature approaches currently exist. In what follows I will consider three promising approaches that have been formulated in recent years: ethical technology assessment, the techno-ethical scenarios approach, and the ETICA approach. For each, I will consider their strengths and weaknesses, after which I will draw a general conclusion.

### Ethical Technology Assessment

Ethical technology assessment (eTA), proposed by Palm and Hansson [7], has as its purpose “to provide indicators of negative ethical implications at an early stage of technological development” (p. 543). Such indicators can subsequently be used to guide design or technology policy. The focus of eTA is on the whole life-cycle of technology development, from initial R&D to ultimate impacts on society. To attain an adequate understanding of future developments, eTA relies on studies in technology assessment (TA)

and on close interactions with developers of technology. The interactions with technology developers are to guarantee an adequate understanding of the technology in question. Studies in TA are to provide insight into both the technology in question and its social consequences, and are also used to organize interactions with technology developers in which eTA is made relevant for the development process. The goal of eTA is not to predict far into the future, but rather to continually assess current practices in technology development and provide feedback to designers and policy makers.

The ethical analysis of an emerging technology takes place by confronting projected features of the technology or projected social consequences with ethical concepts and principles. This yields areas in which a conflict may emerge between the technology and one or more accepted moral principles. This ethical knowledge may then be used to adjust design processes to avoid ethical concerns or to steer decision-making on an emerging technology. Palm and Hansson go on to propose an ethical checklist of nine issues to identify the most common ethical issues in emerging technologies. This list contains issues like privacy, sustainability, issues of control, influence and power and issues of gender, minorities and justice. Not all of these issues are ethical in a conventional sense, but all can be framed as ethical issues.

Palm and Hansson's approach is one of the first ethical approaches explicitly targeted at emerging technologies. It does a good job at advocating the need for ethical TA, and then presents an original approach that seems workable and appears to cover a lot of different issues. Still, the approach has a few limitations. Most importantly, it is rather vague in its methodology, as it does not specify in detail what kind of knowledge needs to be acquired from technology developers and from TA and how it should be acquired, and it also does not spell out in detail how ethical analysis can be performed on the basis of this knowledge. In addition, the ethical checklist of nine items seems somewhat limited, as many recognized moral values and principles are not found on the list, such as autonomy, human dignity, informed consent, distributive justice, and so on. So it would seem one would need a much longer list to be able to do comprehensive ethical assessments of new technologies. Even then, moral issues could be into play for a new technology that are not included in the list. To identify such issues, it would seem that exploring moral

intuitions of either stakeholders or the analyst would be in order.

### The Techno-Ethical Scenarios Approach

The *techno-ethical scenarios approach* of Boenink et al. [2] aims at ethical assessments of emerging technologies that are intended to help policy makers to anticipate ethical controversies regarding emerging technologies. It relies on *scenario analysis*, which is a well-established approach within futures studies. A unique features of the approach is that it aims to anticipate the mutual interaction between technology and morality, and changes in morality that may result from this interaction. Boenink et al. argue that technology may change the way we interpret moral values and may also affect the relative important of particular moral principles. For example, privacy may become a less important principle in an information society where personal information is ubiquitous, and the concept of human responsibility may change in a society in which human decision-making is supported by expert systems. They want to take such changes into account when ethically assessing new technologies, so that new technologies are not evaluated from within a moral system that may not have the same validity by the time an emerging technology has become entrenched in society.

The techno-ethical scenarios approach involves three steps. The first step, "sketching the moral landscape," aims to describe the new technology in question, as well as current moral beliefs, practices and regulations that are directly or indirectly relevant to the technology, and may optionally provide some historical background on the evolution of these beliefs and practices. The second step, "generating potential moral controversies, using NEST-ethics," aims to identify ethical issues and arguments regarding the new technology. This is done using the approach of NEST-ethics [11], which is an approach for identifying ethical issues and arguments in a new technology using a taxonomy of issues and arguments that have been used in past ethical controversies on technology. ("NEST" stands for "New and Emerging Science and Technology".) The NEST-ethics approach performs three tasks. First, it identifies promises and expectations concerning a new technology. Second, it identifies critical objections that may be raised against these promises, for example regarding efficiency and

effectiveness, as well as many conventionally ethical objections, regarding rights, harms and obligations, just distribution, the good life, and others. Third, it identifies chains of arguments and counter-arguments regarding the positive and negative aspects of the technology, which can be used to anticipate how the moral debate on the new technology may develop. During this step, effects of the moral debate on the development of the technology may also be considered. These different steps may involve literature reviews of technologies, promises and expectations, literature reviews of ethical issues, as well as workshops with policy makers and TA experts.

The third step of the techno-ethical scenarios approach, finally, is “constructing closure by judging plausibility of resolutions”. In this step, the multitude of views and arguments from step 2 is reduced by imagining which resolution of the debate is the most plausible. The intention is to use steps 1 through 3 to develop a scenario of how the new technology will develop in the future, how this affects the moral landscape (i.e., moral beliefs, practices and regulations), and how moral closure is eventually reached. The particular scenario they develop, for example, considers how developments in molecular medicine may affect existing moral practices concerning medical experiments with human beings. They project several changes in these practices, based on a scenario study set in Dutch society between 2010 and 2030.

The techno-ethical scenarios approach has some obvious advantages over the eTA approach. It takes into account moral change. It moreover takes on a larger time-frame than eTA, which seems to focus on incremental steps. In addition, it identifies not only ethical issues but also complex patterns of argumentation regarding them. Yet, the techno-ethical scenarios approach has an important limitation as well. This is that it is a descriptive and predictive approach, rather than a normative and prescriptive one. It describes moral issues that are likely to emerge as the technology progresses, not ones that ought to emerge from an ethical point of view, and it considers how these are likely to be resolved, not necessarily how they ought to be resolved.

What this approach may miss, as a result, are ethical issues that are unlikely to collect much public attention but that are nevertheless important. As I have argued in earlier work, important moral controversies may remain hidden because of the complexity or opaqueness of technological artifacts or practices [3]. Such

controversies are not likely to be included in techno-ethical scenarios. Conversely, moral controversies may ensue that are based on a false or misguided understanding of the technology or its social consequences. Such moral controversies do not present moral issues that ought to be considered in assessing emerging technologies, because they are based on false premises. In addition, moral controversies may ensue that are based on parochial moral concerns that would not be considered in an ordinary ethical analysis. My point is hence that moral controversies that may emerge in public debate may be different from moral issues that may result from thorough ethical assessments, even though there may be a large overlap in practice between the two. The current approach focuses on the former type whereas I think an ethical analysis of emerging technology should primarily focus on the latter, as its aim should not be to predict moral debate but to identify normative ethical issues.

### The ETICA Approach

The ETICA approach [9,10] is a recent method for the ethical assessment of emerging information and communication technologies (ICTs).<sup>1</sup> It is so general in scope, however, that nothing prevents its application to other types of technology as well, and it will for this reason be considered as a general approach for the ethical assessment of emerging technology. Thus conceived, the aim of the ETICA approach is to provide comprehensive overviews of ethical issues for emerging technologies that are likely to play out in the medium-term future. The ETICA approach makes use of projections of the future which it derives from futures research. It aims to arrive at a *foresight analysis*, which is a forecasting analysis that considers multiple possible futures, out of which one is chosen as most desirable or important to consider. The ETICA approach relies on multiple futures methods and studies, under the assumption that while individual studies will contain biases and shortcomings, their aggregate use will tend to yield more reliable results.

Ideally the ETICA approach would include doing one’s own future studies, as its researchers say. However, in their study of emerging ICT’s, limitations in resources limit them to two methods for identifying

<sup>1</sup> See also <http://www.etica-project.eu/>, especially the deliverables.



ethical issues in emerging technology. The first is to extract ethical issues from texts about particular emerging technologies in which ethical issues are discussed. Such texts include governmental and political sources, scientific sources such as research reports and journal articles, and non-academic sources such as published future visions of companies. The second is to use bibliometric analysis that finds correlations between emerging technologies and ethical values and concepts in a database of texts on ethics of technology in the academic literature.

The results of multiple futures research studies are used to identify a range of projected artifacts and applications for particular emerging technologies, along with capabilities, constraints and social impacts. These data form the basis for ethical analysis. In the first stage of ethical analysis, the identification stage, ethical issues are identified for particular applications, artifacts or technological properties.<sup>2</sup> Most of the ethical values and principles used in this approach are derived from a prior list of ethical issues for ethical evaluation in a European context. The resulting ethical issues are summarized in a *normative issues matrix*, which specifies relevant normative issues in relation to particular emerging technologies and the artifacts and applications that are expected to result from them. For example, an analysis of robots, as an emerging technology, may focus on particular applications such as service robots in households, robots as companions and robots as soldiers, and discuss ethical aspects of each application. The normative issues matrix also contains more general ethical issues with particular technologies that are not bound to particular applications. For example, an analysis of robots may focus on privacy issues in relation to the sensory capabilities of robots, or responsibility issues in relation to the behavioral autonomy of robots, or ethical issues that are specific to humanoid robots.

At a second stage of ethical analysis, the evaluation stage, the ethical issues of the identification stage are subjected to ethical evaluation and are ranked and ordered in relation to each other. In a third and final stage, the governance stage, governance recommendations are developed for policy makers for dealing with the ethical issues described in the earlier stages.

<sup>2</sup> The ETICA project also uses these data to perform social and legal analyses. However, in my discussion I will focus on its use for ethical analysis.

The ETICA approach is possibly the most elaborate ethical approach to emerging technologies that has been developed to date. It aims at thoroughness by considering a wide range of technological properties, artifacts, applications, and ethical issues. It also engages in ethical evaluation and develops recommendations for governance. And it aims to make use of state-of-the-art work in futures studies. Yet, the approach also has weaknesses. First, its claim to adopt a futures studies approach is somewhat dubious. The main sources of the ETICA approach for locating ethical issues are government and political texts, scientific texts, and non-academic texts. Many of the non-academic and government texts will not be based on scientific methods of futures research. Moreover, many of the scientific texts do not seem to be either. Judging from the literature references in the ETICA projects, many of these texts come from ethics and computer science journals, and most of them do not use methods of futures research.

Second, its assumption that “the overall discourse on future[s] technologies provides as good and reliable an understanding of the future as will be possible to achieve” ([10], p. 9) is also dubious. Rather than merely aggregating predictions about new technologies, it would be better if the approach would provide independent critical assessments of such predictions and the methods used for arriving at them before such predictions are used as a basis for subsequent ethical analysis. It should be granted, though, that in the ETICA project some independent foresight research is undertaken to validate some of the predictions that are made. Third and finally, many of the ethical analysis undertaken in the ETICA project appear to refer to generic properties of the technologies that are studied. In the project these are called “ethical issues stemming from the defining features of the technology” ([6], p. 27). The range of artifacts and implications that is considered is often somewhat limited, and elaborate descriptions of possible artifacts and applications are often missing. For example, in the ethical analysis of robotics, most space goes to the consideration of generic ethical issues, and only a few types of robots and application areas of robotics are considered in detail.

## Conclusion

My review of the three approaches has revealed strong and weak points in each approach. It has also brought

forward various points to consider in an ethics of emerging technologies. A first point is through what approaches and methods technological forecasts are arrived at. The three approaches use various approaches from futures studies and technology assessment, including approaches developed as part of their own approach. A second point concerns the use of ethics and the identification and evaluation of ethical issues. How should this be done? Here, the three approaches also have different answers, though what they have in common is their drive to identify possible ethical issues or controversies and their heuristic use of ethical checklists in doing so. A final point, which has been more implicit in the discussion, concerns the question what an ethics of emerging technology actually studies: is it whole technologies and techniques, is it possible future artifacts, is it uses of artifacts, social consequences, or yet something else? To this question, also, the three approaches give different answers. These three points for an ethics of emerging technologies provide a good challenge to build and improve on the three approaches discussed above. That is what I will turn to in the next section.