

Threats to Autonomy from Emerging ICTs

Brandt Owen Dainow

Maynooth University
Ireland
bd@thinkmetrics.com

Abstract

This paper examines threats to autonomy created by significant emerging ICT's. Emerging ICT's cover a wide range of technologies, from intelligent environments to neuroelectronics, yet human autonomy is potentially threatened by all of them in some way. However, there is no single agreed definition of autonomy. This paper therefore considers the ways in which different versions of autonomy are impacted by different systems. From this range of threats we will derive some properties which any ICT must exhibit in order to threaten human autonomy. Finally, we will show how the range of definitions of autonomy creates problems for customary approaches to value-sensitive design, and indicates a need for greater flexibility when attempting to improve the ethical status of emerging ICT's.

Keywords: ICT; ethics; autonomy; foresight studies; futures research; ETICA; value-sensitive design

1 Introduction

This paper examines ways in which human autonomy can be threatened by emerging ICT's. Emerging ICT's embrace a wide range of technologies, including autonomous systems, intelligent environments, bio-electronic implants, robotics, and artificial intelligence. We will explore the different ways in which each ICT threatens autonomy to show that the nature of the threat, if any, depends on the version of autonomy being used. From this survey of threats we will derive some properties which any ICT must exhibit if it threatens human autonomy, no matter how it is defined. Finally, we will argue the range of definitions of autonomy demonstrates the need for greater flexibility when attempting to improve the ethical status of emerging ICT's.

In order to discuss the complete range of emerging ICT's it is necessary to have a system by which to conceptualise and organise them. We shall use for these purposes the output of the ETICA project (Stahl 2011), which attempted an examination of the complete range of emerging ICT's, developed an ordered taxonomy (Ikonen et al. 2010), and researched the ethical concerns associated with each (Heersmink, van den Hoven, and Timmermans 2010).

2 Methodologies for evaluation of emergent technology

To count as "emerging" (as opposed to "possible") an ICT must be sufficiently developed that we can understand its general features, but which has not yet achieved maturity in all aspects. Since the 1960's the predominant model for such analysis has thus been to treat technologies as "socio-technical systems" (Trist 1981; Lamb and Kling 2003). This position contrasts with the other major treatment of technology, technological determinism, which treats technology as affecting society, but as developing independent of social, cultural, political and other "intangible" causes (Smith and Marx 1994). An *emerging* ICT will have little market penetration (it may still be in prototype), such that its final place and role in the market is yet to become evident; its full and final set of features are unlikely to be completely determined; the business models under which it will function are likely to be nascent; users may still be working out patterns of usage; and the regulatory framework is most probably undeveloped.

Autonomous cars provide a simple example of an emerging ICT which illustrate these uncertainties. Google's autonomous cars are still in prototype (Solveforx 2016), while Tesla vehicles are just beginning to penetrate the market (Statistica 2016). In both cases, the systems

are still in development such that the final set of functions is indeterminate (Solveforx 2016; Musk 2016). The regulatory framework for autonomous vehicles has seen some development in a few places, such as California, but most countries haven't even begun to think about the technology. In the USA, for example, as of March 2017, only eleven states had any regulations regarding this technology, while the majority of attempts to introduce legislation in the USA since 2012 have failed (National Conference of State Legislatures 2017). Business models for the sale and use of this technology are still undetermined. For example, Tesla recently proposed models in which people would be able to send their Tesla car out as an automated taxi when they're not using it (Musk 2016). In the meantime, users are still learning how to use the systems, often in conflict with how the designers intend them to be used. The most dramatic example of this disparity between designer's intent and user behaviour is the frequent driving of Tesla vehicles without human supervision, despite the Tesla's frequent statements that the user should pay attention and keep their hands on the steering wheel at all times (Consumer Reports 2016). Any attempt to understand the ethical impact of autonomous cars must therefore commence with a vision of the state of affairs when the technology has matured – with standardised features, customary patterns of usage, stable regulatory regimes and (reasonably) settled business models. Any attempt to assess the ethical implications of autonomous cars thus becomes an exercise in anticipating what the future will look like. This brings us into the remit of foresight studies.

2.1 Introducing Foresight Studies

'Foresight studies' is one of the terms used to describe the discipline of considering the future in a methodical manner. Foresight studies can be used in any discipline and so the objectives in such research vary widely (Glenn 2009; Godet 2009; Veikko, Kanerva, and Kouri 2009). This presence within many disciplines has led to a variety of terms being used to describe the activity. Common alternative terms for foresight studies are "future studies" (most commonly used in the USA), "futures research" (Europe), "futurology" (Australasia), "prospective studies" (France), "futures field" (Europe) and "prognostics" (Russia) (Veikko, Kanerva, and Kouri 2009). A wide variety of foresight methodologies have evolved and most foresight research projects combine several methodologies.

The different methodologies used within foresight studies can be divided between quantitative and qualitative in technique and between normative and exploratory in purpose, though some methodologies bridge these divides (Veikko, Kanerva, and Kouri 2009; Haegeman et al. 2013; Gordon and Glenn 2004). Quantitative methods are used mainly in predictive forecasting and seek to produce probabilities (Veikko, Kanerva, and Kouri 2009; Haegeman et al. 2013). Quantitative methods are also used in complex modelling, for example, climate modelling. Such methods are based on the assumption that future development is a continuation of past trends. While they provide a continuity from past to future, they cannot allow for disruptive technologies or unexpected developments. Some argue for the need to support quantitative methods with qualitative ones (Veikko, Kanerva, and Kouri 2009; Haegeman et al. 2013) as they are better suited for the anticipation of abrupt or unexpected changes (Ogilvy 2002; Veikko, Kanerva, and Kouri 2009). Foresight research also distinguishes between methodologies on the basis of whether they are normative or exploratory. Normative forecasting is concerned with whether a particular future is desirable and the ethical status of decisions which would lead to it. By contrast, exploratory forecasting explores what is possible, irrespective of its desirability (Veikko, Kanerva, and Kouri 2009). Ethical assessment of the future is rarely a formal component of foresight research project (Bell 1996; Poli 2011; Brey 2012; Clarke 2005; Grunwald 1999; Harris et al. 2010). Key foresight research figures such as Wendell Bell have attempted to determine what ethical values are universally accepted by all peoples in all cultures (Bell 1996). However, Bell has since been criticised for accepting as objective and universal what are merely western cultural values (Poli 2011; Dator 2011; Rubin 2011). A less controversial approach has been to incorporate ethical values as research data. Here distinct sets of exploratory and normative data are gathered. For example, the functionality of future technologies can be drawn the visions of those creating them, after

which philosophical sources can be mined for ethical treatments of these functions. This was the approach taken by the ETICA project.

2.2 ETICA

The most comprehensive foresight research project focused on ethical assessment of emerging ICT's so far has been the EU's ETICA (**E**thical Issues of Emerging **I**CT **A**pplications) project, which ran from 2009 - 2011. The project's aims were to identify "significant emerging ICT's" (Stahl 2011, 1), the ethical issues these gave rise to, and to make recommendations for EU policy makers. A key output from the ETICA project was a taxonomy of emerging ICT's. Here ETICA developed formal structures for technological descriptions and scenario construction (Ikonen et al. 2010). It then used bibliometric analysis to assess published discourse regarding the ethical issues associated with these ICT's (Heersmink, van den Hoven, and Timmermans 2010). The project validated these findings with expert focus groups and surveys (Heersmink et al. 2010). Finally, ETICA developed a series of recommendations (Rainey and Goujon 2011), together with supporting philosophical and methodological expositions (Veikko, Kanerva, and Kouri 2009; Rader et al. 2010; Rainey and Goujon 2011).

ETICA identified 107 different technologies as significant and emerging, which it aggregated into eleven groups. It then identified approximately 400 ethical concerns associated with these technology groups. The aim was to collate the ethical concerns, not explore them in depth, much as an 18th century naval cartographer might map the location and outlines of a chain of islands, but not explore their interior. It is not our intention to review all of these ethical concerns, but to concentrate on issues concerning autonomy. We shall first catalogue how human autonomy can be challenged by each technology group, then elucidate the common characteristics shared by all. However, the meaning of the term 'autonomy' is much contested, so what one considers a threat to autonomy depends on what one means by the term. It is therefore necessary to briefly explore the concept of autonomy. Our aim is not to produce a definition of the term 'autonomy', but to better understand the different ways it may be used when discussing ethical threats.

3 Autonomy

Emmanuel Kant was the first to apply the concept of autonomy to human beings (Schneewind 2007; Dryden 2015; Paul, Miller, and Paul 2003; Dworkin 1988). Before Kant the term 'autonomy' was a purely political one. A state was said to be "autonomous" if its laws were drafted within that state, as opposed to being drafted by a distant imperial court or some similar extra-national body (Dworkin 2015; Schneewind 2007). Kant applied the concept to the individual in *Groundwork of the Metaphysics of Morals* (Kant 1785). His aim was to justify the right of each individual to make their own judgements regarding morality. For the previous 100 years, Enlightenment thinkers had been seeking arguments to combat the politically dominant understanding of ethics, in which ordinary people were seen as too weak-willed to act morally without threats of punishment and promises of reward. Under this view, society was dependant on the guidance of those few exceptional people whom God had enabled to understand and teach His moral laws. The essence of morality for everyone else was to do what they were told (Schneewind 2007). *Groundwork of the Metaphysics of Morals* set out to prove that each person drafted their own rules for their own conduct, that these rules constituted each person's moral code, and that no education was required for this ability - it was universally present in all humans. Kant's labelled the innate capacity of all humans to determine what was morally correct as 'autonomy.' He then used this capacity for autonomy as the basis for human dignity, which then became the basis for human rights. Thus Kant's concept of autonomy created an interlinking of the concepts of individuality, freedom, morality, autonomy, dignity and politics (Schneewind 2007; Bittner 2014; Dryden 2015; Paul, Miller, and Paul 2003).

Since Kant introduced the concept, a number of differing definitions of autonomy have arisen. The range of such accounts, as well as the lack of progress towards any consensus, has led to the concept of "regimes" of autonomy (Anderson 2014; Dworkin 2015) – clusters of similar, but not identical, theories, such that "about the only features held constant from one author to

another are that autonomy is a feature of persons and that it is a desirable quality to have.” (Dworkin 2015, 8). Kant’s criteria for autonomy were, by later standards, rigid and limited. Under Kant’s account, people were required to determine what was right or wrong without reference to the consequences of their actions. Furthermore, under Kant, rules of conduct only counted as a moral if the person believed everyone should act the same way. There was no room for moral pluralism in Kant, or tolerance of other people acting differently (Kant 1998; Guyer 2003; Johnson 2014). This type of definition of autonomy has come to be known as a “substantive procedural” account. A “procedural” account of autonomy takes the position that autonomy is a specific form of thinking process (ie: follows a specific procedure). A subset of procedural accounts are substantive, adding the requirement that autonomy also involves thinking about certain things. By contrast, “content-neutral” procedural accounts try to define autonomy without reference to any particular concern or aim. Content-neutral accounts of autonomy form the majority of modern accounts (Dryden 2015) and are, perhaps, more suited to multi-cultural or pluralistic societies in that they allow individuals to determine for themselves what sources and values to consider when making ethical decisions.

The value given to autonomy varies according to the importance given to competing claims, most frequently those of paternalism and those forms of communitarianism which devalue individual choice in favour of community needs (Bell 2014). Kant’s original concept was developed in an effort to remove external authority as a source of morality. Kant specifically identifies autonomy as oppositional to outside influence. However, we are all influenced by external forces, including family, cultural and religious influences. Consequently, outside influence and the point at which it reduces autonomy is a problem which must be dealt with by every definition of autonomy. One response has been to argue that the concept of autonomy as self-governance isolates the individual from their society and morally devalues family, community, culture and tradition (Christman 2014; Mackenzie and Stoljar 2000; Donchin 2000; Buss and Zalta 2015). These positions have given rise to relational and social conceptions of autonomy based on the necessity of human sociality.

It has also been argued autonomy, if conceived of as only certain forms of thinking, does not account for the full range of human emotional and physical experience, such as trained muscle action (e.g.: playing sport or music) or healthy reflexes, which can also express personal autonomy (Meyers 2005, 1989). These issues collide with particular urgency where ICT’s become concerned with care for the ill or aged (Agich 2003; Burmeister 2016; Dworkin 1988), but also wherever ICT’s come to mediate what someone considers to be core elements of their life.

Kant applied autonomy strictly to the moral realm, though it had political implications. Since then Western Philosophy has come to term the capacity to determine one’s own moral codes as ‘moral autonomy.’ To moral autonomy have been added the concepts of “personal autonomy” (the capacity to determine one’s own actions) and “political autonomy” (the capacity to make one’s own political decisions and have them heeded) (Dryden 2015; Anderson and Christman 2005). What these three definitions share is the concept of “self-governance.” ICT’s threaten human autonomy whenever they interfere with this self-governance. This occurs whenever ICT’s make decisions for people, especially when an ICT imposes on the user a way of doing something which is antithetical to the manner in which the user would have chosen. Some threats to autonomy are universal under all definitions of autonomy, but most depend on the particular form of autonomy being used. Reducing autonomy may be justified on other grounds but such a justification, like the threat, will be based on the version of autonomy being used. It will often be the case that a justification works under one concept of autonomy, but not under others.

4 ETICA’s technology groups and their threats to autonomy

Identification of ethical issues of significant emerging ICT’s is dependent upon first having identified the ICT’s which are both significant and emerging. ETICA defined “significant” technologies as those which were likely to have an important influence on society and

conceptions of the self. It defined “emerging” technologies as those on a stable development path and likely to be in common use by 2030 (Stahl 2011). 107 technologies meeting these criteria were identified, including 3-D printing, sensor networks, augmented reality, location-based services, lab-on-a-chip, molecular electronics, nanobots, mobile payment systems, robotics, social network analysis, speech recognition, internet of things, and wireless power. These were then organised into eleven groups on the basis of shared characteristics, such as type of user, operational context, functionality, and relation to other technologies. For example, both wireless sensor networks and wearable computing were placed into the category of Ambient Intelligence.

The project identified the following eleven technology groups:

1. Affective computing
2. Ambient intelligence
3. Artificial intelligence
4. Bioelectronics
5. Cloud computing
6. Future internet
7. Human/machine symbiosis
8. Neuroelectronics
9. Quantum computing
10. Robotics
11. Virtual and augmented reality

(Ikonen et al. 2010, 42; Stahl 2011, 5)

Not all technology groups were seen by ETICA as generating ethical threats in and of themselves. ETICA assessed quantum computing as unlikely to have any impact on society in the next 10-20 years, while cloud computing and future internet were seen as enablers of other technologies and not as raising unique ethical issues themselves. We shall now briefly examine the remaining technology groups and discuss some of the ways in which each has the potential to reduce human autonomy. Our aim is not to explore the details of the individual threats, but simply to see how autonomy can be threatened by each technology. While a few technologies threaten autonomy no matter how it is defined, the central point of this survey is to demonstrate that most threats only appear under specific definitions of autonomy.

4.1 Affective Computing

Affective computing involves treating human emotion as input and generating output which either emulates human signals of emotion or which seeks to manipulate human emotions. Here ethical concerns derive from the importance of human emotion in communication and the risks of manipulation and misunderstanding.

How affective computing threatens autonomy depends on whether affective computing works or not. Unlike some other technologies, there is no guarantee it will ever be possible to build systems which can accurately read human emotions (Barrett, Gendron, and Huang 2009; Picard 2003). The greatest danger with affective computing lies in the possibility that people will think it is working when it is not. We can usually tell when most technologies operate incorrectly. However, there is no simple way of checking to see if an affective computing system’s assessment of a person’s emotion is accurate. Incorrect assessment of someone’s emotions threaten their autonomy when that data is used to make decisions which affect them.

If ICT’s can successfully understand human emotions, affective computing threatens privacy through the ability to glean information about people which they do not wish revealed. Wherever such information is used to make decisions affecting the person, it may constitute a restriction of their autonomy. Here it depends on the form of autonomy one subscribes to. Some “second-order” procedural accounts of autonomy hold that autonomy is preserved provided one would have agreed with the decision if they had been given the chance (Dworkin 1988, 2015). However, accounts of autonomy which do not allow for such “second-order” assessments hold that any decision denied the person is a restriction on their autonomy. If

those decisions were made in an effort to benefit the person, they *may* be justifiable on the grounds of form of paternalism. Otherwise they simply constitute a form of oppression.

Affective computing therefore constitutes a threat to autonomy whenever affective data is used to make decisions about someone, unless one is using a second-order procedural definition of autonomy *and* the person would have agreed with the decision if they had been given the chance.

Emulation of emotion makes possible threats to autonomy in that it offers the chance to overwhelm rationality with emotion. Any manipulation of a person's emotions which prompts them to make a decision they would not have made otherwise constitutes a restriction of autonomy, irrespective of how autonomy is conceived. The concern is therefore that introducing affective output changes the power balance between system and user. If the ability to accurately detect emotions is combined with the ability to emulate emotional expression, the power of such systems to manipulate and persuade becomes greater than any previous human invention. In this sense, affective systems become human manipulation technology.

4.2 Ambient Intelligence

Ambient intelligence refers to IC technology which is embedded into the environment. Its defining characteristic is the invisibility of the devices, often associated with automated input (for example, a change in room temperature) and less-than-obvious responses by the ambient system (for example, changing the air conditioning settings).

Ambient intelligence offers the capability of personalising the environment to the individual user. Where personalisation is under control of the user, it represents an extension of their autonomy. Where personalisation is outside the user's control, it represents a reduction in autonomy under most, but not all, definitions. Second-order accounts allow for the preservation of autonomy if the user would have agreed to the personalisation had they been given the chance. Only under such accounts is autonomy preserved by automated environmental personalisation. If the personalisation is intended to improve the user's quality of life, reductions in autonomy may be justifiable on the grounds of paternalism. However, that is a problematic debate which cannot be resolved in terms of universal principles, but will ride on the specifics of each case; the type of personalisation, the individuals involved and their personal values.

Ambient intelligence shares several ethical issues with affective computing. Autonomy is threatened by poor quality personalization services, either as a result of an inadequate understanding of human nature, or by the skewing of personalisation protocols resulting from commercial interest or developer ignorance (or bias). Inference of user needs from their behaviour is particularly problematic, especially with regard to children. There is the danger that users will be forced to change their behaviour to get the best out of ambient systems, and so end up being "trained" by their environment's designers (Soraker and Brey 2007).

Ambient intelligence creates the risk of the personal environment acquiring power over people. Since ambient intelligence makes the personal environment controllable by third parties, the personal environment becomes a contestable zone, open to commercial and state interests, both in terms of gathering information and seeking control. Some contestation of the personal environment has always occurred. For example, disputes over noise can be traced back to the earliest cities (Goldsmith 2012). However, ambient intelligence represents such a significant increase in power that contestation for the personal environment becomes a new type of issue – the power to (potentially) totally control someone else's personal space.

4.3 Artificial Intelligence

ETICA examined both "hard" AI (emulation of human thought) and "soft AI" (e.g.: expert systems and software agents). Ethical concerns focused on the use of soft AI as an enabler for other emerging technologies, such as ambient intelligence and affective computing. While issues associated with hard AI were considered, ETICA's judged that hard AI was unlikely to do more than appear in earliest prototypes by 2030 and so was outside the scope of the project.

Our discussions of ambient intelligence and affective computing have both shown how human autonomy can be threatened by the power of these systems over people. AI significantly enhances this power, and so can be an enabler of threats to autonomy wherever it is deployed within ICT systems. Particularly problematic is the possibility that AI may enable systems to learn things about people not anticipated by designer or user. Implementation of informed consent becomes extremely difficult when one does not know what information may be acquired. It has been suggested the solution is to build ethical reasoning into AI systems (Dennis et al. 2013). However, ethical systems are abstract principles of value, not problem-solving algorithms (Kraut 2016; Graham 2004). Ethical dilemmas only exist where multiple ethical values conflict because it is these conflicting values themselves which create the dilemma. AI developers are not, therefore, able to reach for an existing corpus of ethical problem-solving algorithms as if the matter were a simple sort task. Any implementation of ethical processing will therefore impose someone's particular ethical values on the system. Should such an ethical value set not accord with the user's set of ethical values, the user's autonomy will be compromised by the actions of their AI-enabled system. The only way to counter such a threat would be to allow users to customise their AI-enabled systems with their own ethical values and reasoning.

4.4 Bioelectronics

ETICA defined bioelectronics as ICT systems which interact directly with the human body. The primary ethical considerations were focused on use in health care, but those who see bioelectronics as a path to artificial human bodies were also considered. Here bioelectronics represents the potential for technological modification of the human form. It therefore directly confronts the essence of what it is to be human and impacts core human values, such as autonomy, freedom and dignity.

Bioelectronics can be used to change people's internal states by delivering medicines and intervening in bodily processes and thus has the potential to reduce autonomy. However, this concern is dependent on the concept of autonomy used. If one accepts a second-order procedural concept of autonomy (Dworkin 1988), autonomy may be preserved *post-hoc* by agreeing with the earlier bioelectronic intervention. However, conceptions which situate autonomy at the moment of decision must see any bioelectronic intervention as a reduction of autonomy. Whether this loss of autonomy can be justified on paternalistic grounds depends on the intent behind the intervention. However, it may be possible to reframe the issue such that autonomy is not threatened. Discussions of bioelectronics tend to treat intervention as morally equivalent to an intentional act. Under this perspective, each individual bioelectronic intervention is treated as a discrete act to which assent may be given or withheld. This raises problems for the preservation of autonomy because of the requirement for ongoing consent (Agich 2003). However, where this intervention is frequent and automated, such as is the case with drug delivery, it may be more productive to think of it as a form of reflex. Some definitions of autonomy incorporate reflex into their schema as non-cognitive expressions of the self (Meyers 2005, 1989). Treating bioelectronic interventions as artificial reflexes removes any threat to autonomy under such definitions.

The nature of bioelectronic intervention means it is difficult to resist. This places a great deal of potential power in the hands of those controlling the devices; granting them the ability to control someone to a degree not possible by other means. The capability of bioelectronics to change internal bodily states and processes is the same as that of drugs. Accordingly, bioelectronics shares many of the same issues pertaining to autonomy as do drugs. There are today debates about the advisability of treating some mental states as illnesses, concerns about inappropriate use of drugs to control behaviour in elderly and children, and other issues relating to appropriate boundaries to medical intervention (Smith 2012). These same issues apply to interventions by bioelectronic devices. Threats to autonomy apply here in two ways; firstly, "authentic" conceptions of autonomy involve the concept of authenticity (that there is some given essential nature to each individual), such that autonomy requires being in conformance to that essential self (Kühler and Jelinek 2013). These accounts hold that there is

something innate about humans which cannot be changed without harming their authenticity, such as particular forms of cognition (Frankfurt 1971). Theological accounts of autonomy may also depend upon a divinely-ordained authentic human nature (Wilson 1978; Niebuhr 2004). Meanwhile, “coherentist” accounts base autonomy on continuity with personal history, such that changes or decisions which radically depart from someone’s previous patterns or tastes are inauthentic, and thus reductions in autonomy (Miller 1981; Ekstrom 1993). Bioelectronic interventions which take a person away from an essential self, however that self is conceived, are reductions to these forms of autonomy. Finally, where bioelectronic interventions are unwelcome and forced on the individual, they constitute a clear reduction in autonomy under all definitions of autonomy.

4.5 Human-Machine Symbiosis

ETICA characterised human-machine symbiosis as the pairing of innate human capabilities with ICT. This definition covers an extremely wide range of systems, including haptic interfaces, decision-support systems, computer-assisted surgery, augmented reality and direct neural interfaces. There is considerable overlap with other technologies, especially bio- and neuroelectronics and artificial intelligence.

Attempts to enhance human beings through the addition of ICT components may constitute reductions in autonomy under the authenticist views just outlined. Therapeutic devices, designed to replace lost human capabilities rather than enhance them, may be autonomy-preserving, but other accounts can hold that even these constitute a loss of autonomy (Bublitz and Merkel 2009). This is particularly the case where those devices require monitoring or control by others. A device may therefore reduce autonomy when it is operated by someone else, but enhance autonomy where it is controlled by the user (Sharon 2017). Such concerns are not limited to devices implanted into people’s bodies. Use of external devices, for example, as aids to memory, may constitute reductions to autonomy, especially under accounts based on the concept of an “extended self” (Olson 2011; Rachlin and Jones 2010), in which parts of our personal identity are embedded in external objects, such as clothing or mobile phones (Ahuvia 2005; Belk 1988; Turkle 2011). For someone who embeds part of their identity in their phone, even something as simple as an enforced software update may constitute a reduction in their autonomy.

4.6 Neuroelectronics

ETICA defined neuroelectronics as technology interfacing between the human nervous system and electronic devices. While neuroelectronics clearly overlaps with bioelectronics, ETICA felt the intimate connection between the person and their brain meant neuroelectronics could not be adequately examined within a general treatment of bioelectronics.

As with bioelectronics, autonomy is not threatened when neuroelectronics is used to gather data, but only when used to induce changes. Since changes cannot be made without knowledge of internal states, gathering data is, however, a privacy concern as an enabler of reductions to autonomy. Any neuroelectronic system which produces changes within the person constitutes a threat under most accounts of autonomy. While this is clearly the case with conceptions of autonomy which involve an authentic nature or coherent life-history, it may also be the case under second-order procedural accounts (which allow for autonomy preservation if someone agrees to a change afterwards). This is because a person’s agreement to a previous procedure may be merely the result of the changes they have undergone. Where these devices are installed to deal with cognitive impairment, such as dementia, informed consent is impossible and thus autonomy must be reduced (though this may be justifiable on paternalistic grounds). As with bioelectronics and human-machine symbiosis, any accounts of autonomy which involve a divinely-derived authentic human nature will view all neuroelectronic changes as reductions in autonomy.

4.7 Robotics

ETICA defined robots as “machines with motor function that are able to perceive their environment and operate autonomously” (Ikonen et al. 2010, 114). ETICA’s focused on new developments which increase the mobility and intelligence of robotic devices, permitting their deployment into wider areas of society, such as the home and healthcare, and focused on robots built for specific roles within these contexts. ETICA did not consider general-purpose, fully mobile devices controlled by strong artificial intelligence seeking to make war on humanity.

The degree to which robotics can threaten autonomy depends on the definition of autonomy used and the use made of the robot. As has been noted earlier, traditional procedural accounts of autonomy as self-determination have been criticised for unrealistically portraying people as isolated individuals and leaving no space in the account for communitarian elements such as the influence of family or culture (Mackenzie and Stoljar 2000; Stoljar 2015). On this basis some have gone so far as to argue autonomy is an unattainable ideal (Strawson 1994). A more common response has been to develop a treatment of autonomy which includes space for the influence of others, such as the concept of a “social self” (Meyers 2005, 44). The central premise in relational accounts of autonomy is the necessary role of other people in the development of one’s values and the centrality of interpersonal relationships to human existence; that it is not possible to function autonomously without the influence of others (Mackenzie and Stoljar 2000; Stoljar 2015; Donchin 2000). Such conceptions of autonomy are threatened when robots replace humans in roles which have social externalities, such as health care and education. While robots are able to undertake the central tasks, the loss of the human contact constitutes a reduction in relational autonomy. The consequence of this is that there may be some tasks for which robots are not ethically suitable, for no other reason than that they are robots. If the social externalities of a work role are essential for the maintenance of the recipient’s autonomy, then that role must be reserved for humans. Were we to replace nurses or teachers with robots, we could thus see people argue they have an inalienable human right to refuse to be served by a robot and to demand human service.

4.8 Virtual/Augmented Reality

ETICA’s union of virtual reality with augmented reality into one technology group was based on the common feature they share - the imposition of digital output onto the human sensory field. ETICA defined virtual reality as occurring where digital output completely replaced sensory data. Where it did not completely replace external sensory data, ETICA used the term ‘augmented reality’ (Heersmink, van den Hoven, and Timmermans 2010, 114). ETICA’s concerns were founded on the fact that such systems mediate or replace interaction with the physical environment.

Autonomy is threatened when virtual or augmented realities depict objects, people and places in the real world and thus become involved with the many ethical issues associated with depiction, cultural bias and its influence the development of attitudes and taste (Zimbardo and Leippe 1991). The imposition of foreign cultural models in virtual realities would constitute a threat to relational and social conceptions of autonomy (Tomlinson 1991), while lack of alternatives and lack of configuration options threatens all conceptions of autonomy.

4.9 Summary

Many of the technologies examined provide personalised services tailored to the individual. These technologies threaten human autonomy when they fail to personalise effectively. This may result from misreading the user, insufficient granularity within the personalisation, lack of user control, inadequate modelling of the user or by imposing on the user ways of living which are contrary to their values. A number of the technologies change power balances within society, granting for the first time, or greatly increasing, the actant power of the environment over the individual. This power threatens autonomy if used to impose on the user or to enable autonomy-limiting personalisation. Some conceptions of autonomy contain elements which can be disrupted by some ICT’s. Conceptions of autonomy which include inter-personal relationships are challenged by robotics, while conceptions of autonomy which make reference

to authenticity, coherent life-history, or some form of given human nature, are challenged by bio- and neuro-electronics and may also be challenged by human-machine symbiosis.

Irrespective of the nature of autonomy used, our survey of ETICA's research suggests there are certain characteristics any ICT is likely to have if it threatens autonomy:

1. **Surveillance:** Surveillance consists of the obtaining of information pertaining to an individual from their behaviour or communication, followed by the use of that data for purposes, either unknown to that individual or against their wishes (Lyon, Ball, and Haggerty 2012). While it can be used in a paternalistic fashion to enhance the individual's autonomy, our concern is when it is used in a manner which limits the individual's autonomy. The presence of a surveillance system is a necessary precondition for the use of personal information by others and for personalisation services because it constitutes the means by which information supporting personalisation is gathered. Control of the outflowing data by the individual is therefore an effective way in which to retain autonomy. Thus control of privacy is an enabling right to autonomy.
2. **Disparity of Control:** Disparity of control occurs when a third party has a greater control over the use of one's personal data and the autonomy-limiting processes than oneself. This lack of control can be accomplished through lack of knowledge on the part of the individual as to what is occurring, through "take it or leave it" terms of use combined with a lack of alternatives, and through the lack (or concealment) of user configuration capabilities. In some cases, such as data profiling, not just functions, but entire industries may be concealed (Federal Trade Commission 2014; Turow 2011; Dainow 2015).
3. **Insufficient Configurability:** Autonomy can be threatened by lack of variability in configuration options suitable to reflect the variations in the user's desires, style of operation or range of outcomes. This lack of variation can stem from:
 - A lack of recognition by the ICT developers of the need for, or even existence of, such variations.
 - Insufficient consideration of need for user variability during the design process.
 - Insufficient granularity of configuration options.
 - Systemic biases within the delivery mechanisms, including business models, market competition, regulatory framework or any other factors regarding the operational delivery of services. For example, a system might be highly configurable, but delivered to users with a standardised configuration which cannot be changed in an effort to reduce support costs.
4. **Insufficient variation in operational models:** Many new ICT's are owned and operated by a very limited number of providers (Noam et al. 2003; Hillis, Jarrett, and Petit 2013). Often a single provider dominates the market to near-monopoly levels. This limits choice of service provision, and consequently the model under which it is provided. In many cases the business model is identical across all providers within an ICT sector, such that change of provider does not change the circumstances under which the user accesses the service. For example, social networking is only available under a capitalist for-profit model in which service provision is exchanged for personal data which is then commoditised (or monetised). Non-profit and privacy-preserving social network systems do not exist to the degree that a real choice is available. In addition, IP protection promotes walled gardens and suppresses interoperability, further limiting options for choice by users because it is not possible to interoperate between social network providers.

5 Conclusions

The common feature to all conceptions of autonomy is self-determination. What one self-decides, how, on what basis, according to what procedure, and subject to which influences, varies from account to account. Yet efforts to preserve autonomy are central to ICT ethics (Spiekermann 2015). Efforts to reduce the negative ethical impact of ICT's on people through processes such as value-sensitive design need to be cognisant of the various conceptions of autonomy relevant to the intended functionality of the system. This is especially the case with technologies which enter into spheres of life which have previously been purely human or which reduce the power people have over their personal lifeworld. Given the wide range of definitions of autonomy, no single definition can possibly cover all cases.

Because autonomy is frequently central to the ethical status of any technology, it is rarely possible to categorise any ICT as threatening or reducing autonomy without reference to the form of autonomy used. Methodologies to improve the ethical sensitivity of ICT technology, such as value-sensitive design, make frequent reference to the need to specify a particular brand of ethics, such as deontology or utilitarianism, from which to draw values (Spiekermann 2015; Nissenbaum 2001; Winfield, Blum, and Liu 2014), but the application of such values still depends on the concept of autonomy being used. It is generally assumed that there is a single, fixed meaning to autonomy, at least in any given context, and that preservation of autonomy simply requires, at most, identification of the correct version (Manders-Huits 2011). Some have noted the meaning of autonomy can vary with different contexts (Burmeister 2016), but it is possible that it only appears the meaning has changed because the difficulties perceived apply to one version of autonomy and not others, thus making the problematic version of autonomy obvious. Our review of ETICA's survey of threats to autonomy has revealed that many versions of autonomy may be applicable within the one context, some of which may have problems while others do not. Thus, attempts to find the "correct" version of autonomy for any context may be moot. On this basis, it seems more appropriate to develop ICT functionality which does not depend on a single definition of autonomy, but adapts to the user's own definition. This implies a loose pairing of system behaviour to coding, such that output can be fine-tuned after deployment, either by a user, their technical support staff or through the use of soft AI components focused on adapting the system to user feedback.

All significant emerging ICTs can threaten autonomy, but there is no simple way of avoiding this. Efforts such as value-sensitive design stand more chance of success if they focus less on incorporating a single set of values and focus instead on ways by which users may adapt ICT's to reflect their own individual values.

References

- Agich, George J. 2003. *Dependence and Autonomy in Old Age: An Ethical Framework for Long-Term Care*. 2nd ed., Rev. Cambridge, UK ; New York: Cambridge University Press.
- Ahuvia, Aaron C. 2005. 'Beyond the Extended Self: Loved Objects and Consumers' Identity Narratives'. *Journal of Consumer Research* 32 (1): 171–184.
- Anderson, Joel. 2014. 'Regimes of Autonomy'. *Ethical Theory and Moral Practice* 17 (3): 355–68. doi:10.1007/s10677-013-9448-x.
- Anderson, Joel, and John Christman. 2005. 'Introduction'. In *Autonomy and the Challenges of Liberalism: New Essays*, 1–23. Cambridge, UK ; New York: Cambridge University Press.
- Barrett, Lisa Feldman, Maria Gendron, and Yang-Ming Huang. 2009. 'Do Discrete Emotions Exist?' *Philosophical Psychology* 22 (4): 427–37. doi:10.1080/09515080903153634.
- Belk, Russell W. 1988. 'Possessions and the Extended Self'. *Journal of Consumer Research* 15 (2): 139–168.

- Bell, D. 2014. 'Communitarianism'. In *The Stanford Encyclopaedia of Philosophy*, edited by Edward N. Zalta, Summer 2014. <http://plato.stanford.edu/archives/sum2014/entries/communitarianism/>.
- Bell, Wendell. 1996. *Foundations of Future Studies*. New Jersey: Transaction Publishers.
- Bittner, Rüdiger. 2014. 'Autonomy Modest'. *Erkenntnis* 79 (7): 1329–39.
- Brey, Philip A. E. 2012. 'Anticipatory Ethics for Emerging Technologies'. *NanoEthics* 6 (1): 1–13. doi:10.1007/s11569-012-0141-7.
- Bublitz, Jan Christoph, and Reinhard Merkel. 2009. 'Autonomy and Authenticity of Enhanced Personality Traits'. *Bioethics* 23 (6): 360–374.
- Burmeister, Oliver K. 2016. 'The Development of Assistive Dementia Technology That Accounts for the Values of Those Affected by Its Use'. *Ethics and Information Technology*, 1–14. doi:10.1007/s10676-016-9404-2.
- Buss, Sarah, and Edward N. Zalta. 2015. 'Personal Autonomy'. *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/personal-autonomy/>.
- Christman, John. 2014. 'Relational Autonomy and the Social Dynamics of Paternalism'. *Ethical Theory and Moral Practice* 17 (3): 369–82. doi:10.1007/s10677-013-9449-9.
- Clarke, Steve. 2005. 'Future Technologies, Dystopic Futures and the Precautionary Principle'. *Ethics and Information Technology* 7 (3): 121–126.
- Consumer Reports. 2016. 'Tesla's Autopilot: Too Much Autonomy Too Soon'. Consumer Reports. <http://www.consumerreports.org/tesla/tesla-autopilot-too-much-autonomy-too-soon/>.
- Dainow, B. 2015. 'Digital Alienation as the Foundation of Online Privacy Concerns'. *Computers & Society ETHICOMP Special Issue*: 109–17. doi:10.1145/2874239.2874255.
- Dator, Jim. 2011. 'Wendell Bell: The Futurist Who Would Put My Grandmother in Prison'. *Futures* 43 (6): 578–582.
- Dennis, Louise, Michael Fisher, Marija Slavkovic, and Matt Webster. 2013. 'Ethical Choice in Unforeseen Circumstances'. In *Towards Autonomous Robotic Systems*, 433–445. Springer.
- Donchin, Anne. 2000. 'Autonomy and Interdependence: Quandaries in Genetic Decision-Making'. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 236–58. New York: Oxford University Press.
- Dryden, Jane. 2015. 'Autonomy'. *Internet Encyclopedia of Philosophy*. <http://www.iep.utm.edu/autonomy/>.
- Dworkin, Gerald. 1988. *The Theory and Practice of Autonomy*. Cambridge Studies in Philosophy. Cambridge ; New York: Cambridge University Press.
- . 2015. 'The Nature of Autonomy'. *Nordic Journal of Studies in Educational Policy* 1 (0). doi:10.3402/nstep.v1.28479.
- Ekstrom, Laura Waddell. 1993. 'A Coherence Theory of Autonomy'. *Philosophy and Phenomenological Research* 53 (3): 599–616.
- Federal Trade Commission. 2014. 'Data Brokers: A Call for Transparency'. Federal Trade Commission. <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>.
- Frankfurt, Harry G. 1971. 'Freedom of the Will and the Concept of a Person'. *The Journal of Philosophy* 68 (1): 5–20.

- Glenn, Jerome C. 2009. 'Introduction to Futures Research Methods'. In *Futures Research Methods*, edited by Jerome C Glenn and Theodore J Gordon, 3rd ed. Washington, DC: The Millenium Project.
- Godet, Michel. 2009. '11. Structural Analysis'. In *Futures Research Methods*, 3rd ed. Washington, DC: The Millenium Project.
- Goldsmith, Mike. 2012. *Discord: The Story of Noise*. Oxford: Oxford University Press.
- Gordon, Theodore J, and Jerome C Glenn. 2004. "Integration, Comparisons, and Frontiers of Futures Research Methods". In *EU-US Seminar: New Technology Foresight, Forecasting & Assessment Methods*. Vol. 1314.
- Graham, Gordon. 2004. *Eight Theories of Ethics*. London; New York: Routledge/Taylor and Francis Group.
- Grunwald, Armin. 1999. 'Technology Assessment or Ethics of Technology? Reflections on Technology Development between Social Sciences and Philosophy.' *Ethical Perspect* 6 (2): 171 – 182.
- Guyer, Paul. 2003. 'Kant on the Theory and Practice of Autonomy'. In *Autonomy*, edited by Ellen Frankel Paul, Fred Dycus Miller, and Jeffrey Paul, 70–98. Cambridge: Cambridge University Press.
- Haegeman, Karel, Elisabetta Marinelli, Fabiana Scapolo, Andrea Ricci, and Alexander Sokolov. 2013. 'Quantitative and Qualitative Approaches in Future-Oriented Technology Analysis (FTA): From Combination to Integration?' *Technological Forecasting and Social Change* 80 (3): 386–397.
- Harris, Ian, Penny Duquenoy, Richard Jennings, David Pullinger, and Simon Rogerson. 2010. 'Ethical Assessment of New Technologies: A Meta-Methodology'. *Journal of Information, Communication and Ethics in Society* 9 (1): 49–64.
- Heersmink, Richard, Jeroen van den Hoven, and Job Timmermans. 2010. 'D.2.2 Normative Issues Report'. D.2.2. ETICA Project.
- Heersmink, Richard, Job Timmermans, Jeroen van den Hoven, and Kutoma Wakunuma. 2010. 'D.2.3 Normative Issues Workshop Report'. D.2.3. ETICA Project.
- Hillis, Ken, Kylie Jarrett, and Michael Petit. 2013. *Google and the Culture of Search*. New York: Routledge.
- Ikonen, Veikko, Minni Kanerva, Panu Kouri, Bernd Stahl, and Kutoma Wakunuma. 2010. 'D.1.2. Emerging Technologies Report'. D.1.2. ETICA Project.
- Johnson, Robert. 2014. 'Kant's Moral Philosophy'. Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/sum2014/entries/kant-moral/>.
- Kant, Immanuel. 1785. *Grundlegung Zur Metaphysik Der Sitten*. Riga: Johann Friedrich Hartnoch.
- . 1998. *Groundwork of the Metaphysics of Morals*. Translated by Mary J. Gregor. Cambridge Texts in the History of Philosophy. New York: Cambridge University Press.
- Kraut, Richard. 2016. 'Aristotle's Ethics'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2016. <http://plato.stanford.edu/archives/spr2016/entries/aristotle-ethics/>.
- Kühler, Michael, and Nadja Jelinek, eds. 2013. *Autonomy and the Self*. Dordrecht: Springer Netherlands. <http://link.springer.com/10.1007/978-94-007-4789-0>.
- Lamb, Roberta, and Rob Kling. 2003. 'Reconceptualizing Users as Social Actors in Information Systems Research'. *MIS Quarterly*, 197–236.

- Lyon, David, Kirstie Ball, and Kevin D Haggerty. 2012. *Routledge Handbook of Surveillance Studies*. Abingdon, Oxon; New York, NY: Routledge.
- Mackenzie, Catriona, and Natalie Stoljar, eds. 2000. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. New York: Oxford University Press.
- Manders-Huits, Noëmi. 2011. 'What Values in Design? The Challenge of Incorporating Moral Values into Design'. *Science and Engineering Ethics* 17 (2): 271–287. doi:10.1007/s11948-010-9198-2.
- Meyers, Diana T. 1989. *Self, Society, and Personal Choice*. New York: Columbia University Press.
- . 2005. 'Decentralizing Autonomy: Five Faces of Selfhood'. In *Autonomy and the Challenges to Liberalism*, edited by John Christman and Joel Anderson, 27–55. Cambridge: Cambridge University Press. <http://ebooks.cambridge.org/ref/id/CBO9780511610325A011>.
- Miller, Bruce L. 1981. 'Autonomy & the Refusal of Lifesaving Treatment'. *Hastings Center Report* 11 (4): 22–28.
- Musk, Elon. 2016. 'Tesla Master Plan, Part Deux'. *Personal Blog*. July 20. <https://www.tesla.com/blog/master-plan-part-deux>.
- National Conference of State Legislatures. 2017. 'Autonomous | Self-Driving Vehicles Legislation'. National Conference of State Legislatures. <http://www.ncsl.org/research/transportation/autonomous-vehicles-legislation.aspx>.
- Niebuhr, R. 2004. *The Nature and Destiny of Man: A Christian Interpretation: Human Nature*. Library of Theological Ethics. Westminster, John Knox Press.
- Nissenbaum, H. 2001. 'How Computer Systems Embody Values'. *Computer* 34 (3): 118–20. doi:10.1109/2.910905.
- Noam, Eli M, Donald Hay, Michael R Baye, and John Morgan. 2003. 'The Internet: Still Wide Open and Competitive?' *OII Internet Issue Brief*, no. 1.
- Ogilvy, James. 2002. 'Futures Studies and the Human Sciences: The Case for Normative Scenarios'. In *New Thinking for a New Millennium: The Knowledge Base of Futures Studies*, edited by Richard Slaughter. London: Routledge.
- Olson, Eric T. 2011. 'The Extended Self'. *Minds and Machines* 21 (4): 481–495.
- Paul, Ellen Frankel, Fred Dycus Miller, and Jeffrey Paul, eds. 2003. *Autonomy*. Cambridge: Cambridge University Press.
- Picard, Rosalind W. 2003. 'Affective Computing: Challenges'. *International Journal of Human-Computer Studies* 59 (1): 55–64.
- Poli, Roberto. 2011. 'Ethics and Future Studies'. *Int. J. Management Concepts and Philosophy* 5 (4): 403–10.
- Rachlin, Howard, and Bryan A. Jones. 2010. 'The Extended Self.' In *Impulsivity: The Behavioral and Neurological Science of Discounting.*, edited by Gregory J. Madden and Warren K. Bickel, 411–31. Washington: American Psychological Association. <http://content.apa.org/books/12069-015>.
- Rader, Michael, A. Antener, Rafael Capurro, Michael Nagenborg, and L. Stengel. 2010. 'D.3.2 Evaluation Report'. ETICA Project.
- Rainey, Stephen, and Philippe Goujon. 2011. 'D.4.2 Governance Recommendations'. D.4.2. ETICA Project. <http://www.academia.edu/download/30848560/Deliverable4.2.pdf>.

- Rubin, Anita. 2011. 'Living in the Age of Emotional Rationality: Wendell Bell, Social Media and the Challenges of Value Change'. *Futures* 43 (6): 583–589.
- Schneewind, Jerome B. 2007. *The Invention of Autonomy: A History of Modern Moral Philosophy*. 7th ed. Cambridge: Cambridge Univ. Press.
- Sharon, Tamar. 2017. 'Self-Tracking for Health and the Quantified Self: Re-Articulating Autonomy, Solidarity, and Authenticity in an Age of Personalized Healthcare'. *Philosophy & Technology* 30 (1): 93–121. doi:10.1007/s13347-016-0215-5.
- Smith, Brendan L. 2012. 'Inappropriate Prescribing'. *Monitor on Psychology* 43 (6): 36.
- Smith, Merritt Roe, and Leo Marx. 1994. *Does Technology Drive History?: The Dilemma of Technological Determinism*. Cambridge, Mass.: MIT Press.
- Solveforx. 2016. 'Google Self-Driving Car Project Website'. *Google Self-Driving Car Project*. <https://www.google.com/selfdrivingcar>.
- Soraker, Johnny Hartz, and Philip Brey. 2007. 'Ambient Intelligence and Problems with Inferring Desires from Behaviour'. *International Review of Information Ethics* 8: 7–12.
- Spiekermann, Sarah. 2015. *Ethical IT Innovation: A Value-Based System Design Approach*. Boca Raton, Fl.: Taylor & Francis.
- Stahl, Bernd Carsten. 2011. 'Project Final Report'. ETICA Project.
- Statistica. 2016. 'Statistics and Facts about Tesla'. *The Statistics Portal*. <http://www.statista.com/topics/2086/tesla/>.
- Stoljar, Natalie. 2015. 'Feminist Perspectives on Autonomy'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2015. <http://plato.stanford.edu/archives/fall2015/entries/feminism-autonomy/>.
- Strawson, Galen. 1994. 'The Impossibility of Moral Responsibility'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 75 (1/2): 5–24.
- Tomlinson, John. 1991. *Cultural Imperialism: A Critical Introduction*. Baltimore, Md.: Johns Hopkins University Press.
- Trist, Eric L. 1981. *On Socio-Technical Systems*. Ontario: Ontario Ministry of Labour.
- Turkle, Sherry. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books. <http://public.eblib.com/choice/publicfullrecord.aspx?p=684281>.
- Turow, Joseph. 2011. *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your World*. New Haven: Yale University Press.
- Veikko, Ikonen, Minni Kanerva, and Panu Kouri. 2009. 'D.1.1 Heuristics & Methodology Report - Final'. D.1.1. ETICA Project.
- Wilson, Edward O. 1978. *On Human Nature*. Cambridge, Mass.: Harvard University Press.
- Winfield, Alan FT, Christian Blum, and Wenguo Liu. 2014. 'Towards an Ethical Robot: Internal Models, Consequences and Ethical Action Selection'. In *Advances in Autonomous Robotics Systems*, 85–96. Springer.
- Zimbardo, Philip G, and Michael R Leippe. 1991. *The Psychology of Attitude Change and Social Influence*. Philadelphia: Temple University Press.

commercial use, distribution, and reproduction in any medium, provided the original author and AJIS are credited.

