

This is a preprint version of the following article:  
Brey, P. (2000). 'Disclosive Computer Ethics: The Exposure and Evaluation of Embedded Normativity in Computer Technology,' *Computers and Society*, 30(4): 10-16.

# **Disclosive Computer Ethics: The Exposure and Evaluation of Embedded Normativity in Computer Technology**

## **Abstract**

This essay provides a critique of mainstream computer ethics and argues for the importance of an additional approach called disclosive computer ethics, which is concerned with the moral deciphering of embedded values and norms in computer systems, applications and practices. Also, four key values are proposed as starting points for disclosive studies in computer ethics: justice, autonomy, democracy and privacy. Finally, it is argued that research in disclosive computer ethics should be multi-level and interdisciplinary, distinguishing between a disclosure level, a theoretical level, and an application level.

**Keywords:** Disclosive computer ethics; method; value-sensitive design; ethics of design; interdisciplinary research; applied ethics

## 1. LIMITATIONS OF MAINSTREAM COMPUTER ETHICS

The aim of this essay is to outline a particular approach to computer ethics called *disclosive computer ethics*. Although already increasingly influential within computer ethics, disclosive computer ethics still deviates from mainstream approaches in computer ethics. *Mainstream computer ethics* is a name that will be used in this essay for those approaches that currently make up the mainstream in computer ethics. These are approaches that follow what I will call the *standard model of applied ethics*. In this section, I will sketch the standard model of applied ethics, illustrate how mainstream computer ethics fits this model, and then go on to outline limitations of this model that are overcome by disclosive computer ethics. My aim is not to present disclosive computer ethics as a rival approach, but rather as an approach that is complementary to mainstream computer ethics. The remaining sections of the essay will be devoted to an outline and defense of disclosive computer ethics as an approach within computer ethics.

Mainstream computer ethics, I will argue, takes as its point of departure a particular model of applied ethics that may be called the standard model, because it is used in the vast majority of work in applied ethics. Studies in applied ethics that adopt the standard model aim to clarify and evaluate morally controversial practices through an application and defense of moral principles. Research within this model usually proceeds in three steps. First, an individual or collective practice is outlined that has been the topic of moral controversy. A biomedical ethicist, for example, may write about euthanasia, cloning, informed consent, or late-term abortion. Likewise, environmental ethicists may focus on wetlands development, the use of nuclear energy or emission of greenhouse gases. Next, an attempt is usually made to clarify and situate the practice, through conceptual analysis and fact-finding. For example, a moral study of late-term abortion may analyze the concept of personhood, distinguish different stages in the development of the fetus, and describe existing procedures and techniques for late-term abortions. Third, moral principles are outlined, along with moral judgments or intuitions, and applied to the topic. This is a deliberative process, with arguments for and against particular principles or their applicability to the case. The outcome is a moral evaluation of the practice that is investigated.

The standard model of applied ethics is adopted in most of the literature in computer ethics. In mainstream computer ethics, a typical study begins by identifying a morally controversial practice, like software theft, hacking, electronic monitoring, or Internet pornography. Next, the practice is described and analyzed in descriptive terms, and finally, moral principles and judgments are applied to it and moral deliberation takes place, resulting in a moral evaluation, and optionally, a set of policy recommendations.

Three features of mainstream computer ethics are noteworthy in the context of this essay. First, mainstream computer ethics focuses on existing moral controversies. Second, its focus is on practices, i.e. the individual or collective behavior of persons. It aims to evaluate and devise policies for these practices. And third, its focus usually is on the *use* of computer technology, as opposed to, e.g., its design or advertisement. As Moor summed it up in his influential essay ‘What is computer ethics?’ (1985), ‘A typical problem in computer ethics arises because there is a policy vacuum about how computer technology should be used.’ (p. 266)

Here, a first limitation of mainstream computer ethics may be identified. Mainstream computer ethics limit itself to the analysis of morally controversial practices for which a policy vacuum currently exists. But what about computer-related practices that are not (yet) morally controversial, but that nevertheless have moral import? Surely, one would not want such practices to be ignored. Let us call a practice that has moral import but that is not generally recognized as morally controversial a (*morally*) *nontransparent* or *opaque* practice. Clearly, some of the research effort in computer ethics should be devoted to identifying and studying morally nontransparent practices. I take this to be part of the *critical* function of computer ethics.

Computer-related practices are often morally nontransparent because they are unfamiliar or unknown to most people. Most Internet users, for example, are unfamiliar with the ways in which their behavior or personal data may be monitored on-line. So even though on-line monitoring is not an issue for most Internet users, it is part of the critical function of computer ethics to identify, analyze, morally evaluate and devise policy guidelines for on-line monitoring.

Nontransparency may also arise for practices that are familiar in their basic form, but that are not recognized as having the moral implications that they in fact have. The hardware, software, techniques and procedures used in computing practice often has the appearance of *moral neutrality* when in fact it is not morally neutral. For example, search engines used on the Internet, seem to have the innocuous task of helping users to quickly find relevant information. However, as Introna & Nissenbaum (1998) have argued, the particular search algorithms used in search engines are far from neutral, and are often discriminatory, in this way threatening the idea of the Web as a public space, in which there exists for everyone an equal opportunity in letting one’s voice heard.

The above remarks about the moral non-neutrality of many technologies and techniques point to a second limitation of mainstream computer ethics. Mainstream computer ethics focuses on the morality of practices, particularly on the use of computer technology. What is often marginalized in the discussion, or left out entirely, is the moral role of the technology that is being used. That is, the particular features of computer systems and software is often taken as a given in computer ethics. The technology is taken as a neutral tool with which both moral and immoral actions can be performed, and the focus is on the actions that are performed with it. In philosophical

and empirical studies of technology, however, it is by now accepted that technologies are not neutral, and that they often play an active part in shaping their environments. A corollary of this view is that technological artifacts may themselves become the object of moral scrutiny, independently from, and prior to, particular ways of using them.

The notion that technology can have moral properties is an extension of the notion that it can have political properties (e.g., Winner 1980; Sclove 1995; Feenberg, 1999). As Winner (1980) has argued, technological artifacts and systems function much like laws, by constraining behavior and serving as frameworks for public order. Richard Sclove has made the same point by identifying technical artifacts as elements of social structure. Sclove defines the social structure of a society as its 'background features that help define or regulate patterns of human interaction. Familiar examples include laws, dominant political and economic institutions, and systems of cultural belief.' (1995, p. 11). He argues that technologies should also be included in this list, because they have the same kinds of structural effects as these other elements of social structure.

Technologies are, for example, capable of coercing individuals to behave in certain ways, may provide opportunities and constraints, may affect cultural belief systems, and may require certain background conditions for them to function properly.<sup>1</sup> Many such structural effects of technology may be analyzed from a moral point of view. Feenberg, finally, has explained that the political properties of technical artifacts often become opaque because the artifact is accepted in society as apolitical: it is understood in terms of a *technical code*, which is a neutral specification of its technical properties and functions that obscures the nonneutral social and political properties of the artifact.

The notion that technologies are themselves worthy objects of moral analysis is perhaps especially true for computer technology. This is because computer systems have become, because of their information processing abilities, important determinants in many human decision-making processes, behaviors and social relations. Computer systems often function less as background technologies and more as active constituents in the shaping of society. This active role of computer systems warrants that special attention is paid in computer ethics to their design features, as an object of moral analysis largely independently of their use. To conclude, then, mainstream computer ethics has two important, interrelated limitations: it tends to focus too narrowly on publicly recognized moral dilemma's, and it tends to downplay computer technology itself as an object of moral analysis.

---

<sup>1</sup> Extended discussion and examples of these properties of technologies are found in Winner (1980), Sclove (1995), Pfaffenberger (1992) and Akrich (1992).

## **2. HIDDEN MORALITY AND DISCLOSIVE COMPUTER ETHICS**

Disclosive computer ethics is the name I propose for a family of recent approaches in computer ethics that are centrally concerned with the moral deciphering of computer technology. This work distinguishes itself from mainstream computer ethics on exactly the two points mentioned: it tends to be concerned with the uncovering of moral issues and features in computing that had not until then gained much recognition, and its focus tends to be on the design features of computer technology. Disclosive studies in computer ethics are hence studies concerned with disclosing and evaluating embedded normativity in computer systems, applications and practices. Its major contribution to computer ethics is not so much found in the development or application of ethical theory, but rather in the description of computer technology and related practices in a way that reveals their moral importance.

Existing work in computer ethics that takes a disclosive approach covers moral issues such as privacy, democracy, distributive justice, and autonomy, and covers the full spectrum of information and communication technologies. Friedman & Nissenbaum (1997), for example, is a study of bias in computer systems. Biases in computer systems are usually not recognized, but Friedman & Nissenbaum try to reveal the existence of bias by describing computer systems with bias and by bringing into view the possible unjust consequences of such systems (see also Brey, 1998). Similarly, Brey (1999; 1998) is concerned with the consequences of the design of computer systems for the autonomy of users. A large part of the research in these papers is concerned with revealing the potential impacts of computer designs on the autonomy of users, and much less attention is devoted to theorizing and applying moral principles of autonomy. Other examples of disclosive studies in computer ethics are Nissenbaum (1997), who reveals the moral importance of practices of registering public information, Blanchette (1998), who reveals the importance of trust relations in cryptographic protocols and payment mechanisms, Introna & Nissenbaum (1998), who decipher the hidden politics of search engines, Agre and Mailloux (1997), who reveal the implications for privacy of Intelligent Vehicle-Highway Systems, and Johnson (1997), who analyzes antidemocratic tendencies inherent within the structure and organization of the Internet.

Admittedly, the description of technologies and practices so as to reveal their moral importance presupposes that one can already discern what is and what is not

morally important, and hence that relevant moral values have already been formulated before analysis comes off the ground. However, this does not mean that one must already be equipped with moral theories before disclosive analysis can take place. The (potential) moral importance of designs or practices is already sufficiently established if it is shown that these designs or practices yield, for example, an unequal distribution of power or of goods, that they diminish privacy or freedom (according to common-sense notions of these terms), that they affect social relations or statuses, or that they touch on other important moral values that are widely shared in society. Therefore, a more precise moral evaluation can wait until after disclosive analysis.

Thus, a disclosive study in computer ethics may take the form of a two-stage process. In the first stage of analysis, some technology ( $X$ ) is analyzed from the point of view of a relevant moral value ( $Y$ ) (where  $Y$  is, e.g., privacy, justice, freedom, etc.), which is only given a loose, common-sense definition. This analysis may yield a tentative conclusion that certain features of  $X$  tend to undermine (or perhaps sustain)  $Y$  in particular ways. For example, it may be found that search engines in use on the Internet tend to undermine informational privacy, where informational privacy is defined loosely as the control that individuals have over the disclosure of information about their person. This analysis may prompt a second stage in which theories of informational privacy are applied and perhaps further developed so as to arrive at a more specific normative evaluation of the privacy-aspects of search engines, that can also be used to arrive at policy guidelines regarding their design, use and regulation.

Of course, it is also possible to do disclosive analysis in a more theory-driven way. In the above example, one would then start with a moral theory of informational privacy that would contain specific moral principles, and then analyze the manner in which search engines uphold or fail to uphold these principles. Optionally, this analysis could again result in a set of policy recommendations regarding the privacy-aspects of search engines.

These two approaches are both acceptable varieties of disclosive computer ethics. There are, however, at least two reasons why a theory-driven variety may ultimately be less preferable. First, a theory-driven approach tends to make the acceptance of a disclosive analysis dependent on the acceptance of a particular moral theory. For example, a study that shows that existing search engines violate a particular conception of informational privacy found in theory  $T$  may not convince someone that search engines raise issues for informational privacy if that person rejects  $T$ . That person

might have been convinced by an analysis that had started with a loose definition of informational privacy, and proceeded to show that search engines pose a problem for informational privacy according to this loose definition. Second, a theory-driven approach will already contain preconceptions about the technology or practice that is under scrutiny, because it already employs a highly theoretical vocabulary in the analysis of observable phenomena, that may include empirical presuppositions. It may therefore come to observations that are as based in part on preconceptions in the theory that is applied, at points where more neutral descriptions may be preferable. In conclusion, there are good reasons not to choose a theory-driven approach in disclosive computer ethics if given the choice.

### **3. KEY VALUES AS DEPARTURE POINT FOR ANALYSIS**

Disclosive computer ethics hence uncovers and morally evaluates values and norms embedded in the design and application of computer systems. For it to qualify as computer ethics, the values and norms in question must be *moral* values and norms. Many values and norms are nonmoral, including values like efficiency and profit or norms that prescribe the correct usage of words like 'telephone' or the maximum voltage. Although any distinction between moral and nonmoral values and norms is bound to be controversial, moral norms and values are usually recognized as pertaining to the rightness or wrongness, or goodness or badness, of actions and those who perform them, and correspondingly, to the praiseworthiness or blameworthiness of these actions and persons, and often point back to more fundamental values that are thought to jointly make up human conceptions of the Good.

I will proceed to propose four fundamental values that have consistently shown to be important in disclosive studies in computer ethics, and that may therefore also constitute a point of departure for future studies. These are values that are part of the shared moral vocabulary of society, and that have near-universal acceptance, even though very different interpretations exist of them. They are: justice (and fairness), autonomy (and freedom), democracy, and privacy. In what follows, I will propose relatively informal definitions of these concepts and then outline different issues in computer ethics in which they have been shown to play a role.

*(1) Justice (and fairness)*

The notion of justice is usually understood as implying that individuals should not be advantaged or disadvantaged unfairly or undeservedly. This implies, amongst other things, that society should not promote the unfair distribution of social goods. In particular, society should not promote the unfair distribution of very basic social goods, such as rights and liberties, powers and opportunities, income and wealth, and the social bases of self-respect, as these are essential for individuals to carrying out their life plans. Following Rawls (1971), ethicists often call such social goods *primary* social goods. So there is a shared agreement of society that the distribution of primary social goods in society should not unfairly disadvantage some members. Disagreements exist mostly just on particular distribution models of primary social goods.

Disclosive studies of computer systems and justice are studies of ways in which these systems, or particular applications of them, affect unequal distribution of (primary) social goods for which no immediate justification exists. These studies will normally focus on particular social goods, such as powers, freedoms, job opportunities, or social statuses, and relate these to specific types of computer systems and applications, and different social groups. To these analyses, explicit theories of distributive justice may then be applied to argue whether resulting inequalities are indeed unjust and to make policy recommendations.

*(2) Autonomy (and freedom)*

Whereas the notion of freedom is probably familiar to a larger segment of society than the notion of autonomy, the latter notion is arguably more fundamental. Individual autonomy is commonly taken to mean that individuals have a number of rights to individual freedoms, such as the right to freedom of speech, to religion, to peaceful assembly, and to privacy. Freedom rights such as these are fundamental because they ensure that human beings are able to draw out their own life plans that reflect, as much as possible, values and needs of their own, instead of those of the government or of other citizens. In other words, freedom rights protect goods that are fundamental for carrying out one's own life plan. If one has no privacy, if one cannot practice one's religion, or if one cannot speak freely, one lacks some of the most basic goods that are prerequisite to carrying out one's life plan.

The ideal of individual autonomy has long been defended as fundamental to human flourishing and self-development (e.g., Dworkin, 1988; Hill, 1991). Individual

autonomy is often defined as self-governance, that is, the ability to construct one's own goals and values, and to have the freedom to make choices and plans and act in ways that are believed by one to help achieve these goals and promote these values. Individual autonomy is often defended as important because such self-governance is required for self-realization, that is, it is a freedom one must have in order to create a life that is experienced by oneself as meaningful and fulfilling. As Dworkin has pointed out, moreover, individual autonomy may be a requirement for a conception of human beings as equals. If some human beings are not autonomous, they cannot give equal input into moral principles that are aimed to reflect individual preferences, and hence they cannot function as equals in moral life (1988: 30-31).

Disclosive studies of computer systems and autonomy are studies of ways in which these systems or uses of them affect the autonomy or freedom of their users and possible third parties. This includes studies of how computer systems may constrain their users and may help install dependencies (on computer systems themselves, system operators, managers, etc.). It also includes studies of constraints on freedom of information that consider the freedom of individuals to acquire, possess or distribute certain sorts of information. Also, studies of computer systems and autonomy need not be limited to individual autonomy. Autonomy of organizations and nation states, also called independence or sovereignty, is also a widely shared moral goal, and implications of computer technology for sovereignty is therefore also a worthy topic of investigation.

### *(3) Democracy*

Although different conceptions of democracy exist, the idea is widely shared that democracy is rule by 'the people,' and implies that a democratic society or organization has procedures for political decision-making in which every member has the opportunity to exert influence so as to have his or her interests taken into account. This implies that political power is not concentrated in the hands of individuals or organizations that cannot be held accountable by the public but is instead distributed to a greater or lesser degree over all members of society.

Although the value of democracy relates to the values of justice and autonomy, it is generally held to be a separate value, and therefore separate disclosive studies of computer systems and democracy are warranted (cf. Selove, 1995). These are studies designed to investigate whether particular computer technologies or applications effect

redistributions of political power in societies or organizations such that a loss or gain in democratic content is implied. Particularly important is the question who has control over information and communication processes in societies and organizations, as ownership of and control over information has become an important source of political power in society.

#### (4) *Privacy*

Privacy is often taken to be an aspect of autonomy. Yet, because of its importance in relation to computer technology, it merits to be treated separately. Privacy is the freedom granted to individuals to control their exposure to others. A customary distinction is that between *relational* and *informational* privacy. Relational privacy is the control over one's person and one's personal environment, and concerns the freedom to be left alone without observation or interference by others. Informational privacy is one's control over personal information, in the form of text, pictures, recordings, etc. Disclosive studies of computer systems and privacy analyze the implications of computer systems and applications for relational and informational privacy.

Evidently, the above four values do not exhaust the set of moral values that may be taken as the point of departure of disclosive studies in computer ethics. Other values may also be taken as departure points, including values on which not societal consensus exists (e.g., specific religious or cultural values). It may be worth investigating, in particular, to what extent computer technology affects general aspects of the quality of life or quality of society that are often considered valuable (Brey, 1997).

If disclosive studies of computer technology are to qualify as studies in ethics, then it is necessary that analysis be explicitly centered around moral values such as the ones discussed above. A mere investigation of the way in which computer technology affects distributions of wealth and power is not in itself a study in ethics, even if it is motivated by an interest in moral issues, as an explicit normative thematization is lacking. However, such studies may constitute useful inputs for research in computer ethics in which normative issues *are* explicitly thematized.

#### 4. THE NEED FOR MULTI-LEVEL INTERDISCIPLINARY RESEARCH

Disclosive computer ethics requires an approach that is multi-level and interdisciplinary. It is multi-level in that research is to take place at various stages or levels. Three such levels can be discerned. First, there is the *disclosure level*, which is the initial level at which disclosive computer ethics research takes place. At this level, some type of computer system or software is analyzed from the point of view of a relevant moral value like privacy or justice.

Second, there is the *theoretical level*, which is the level at which moral theory is developed and refined. As Jim Moor (1985) has pointed out, the changing settings and practices that emerge with new computer technology may yield new values, as well as require the reconsideration of old values. There may also be new moral dilemmas because of conflicting values that suddenly clash when brought together in new settings and practices. It may then be found that existing moral theory has not adequately theorized these values and value conflicts. Privacy, for example, is now recognized by many computer ethicists as requiring more attention than it has previously received in moral theory. In part this is due to reconceptualizations of the private and public sphere brought about by the use of computer technology, which has resulted in inadequacies in existing moral theory about privacy. It is therefore fitting for computer ethicists to contribute to the development of moral theory about privacy. In general, it is part of the task of computer ethics to *further develop and modify existing moral theory* when existing theory is insufficient or inadequate in light of new demands generated by new practices involving computer technology.

Third, there is the *application level*, in which, in varying degrees of specificity and concreteness, moral theory is applied to analyses that are the outcome of research at the disclosure level. For example, the question of what amount of protection should be granted to software developers against the copying of their programs may be answered by applying consequentialist or natural law theories of property, and the question of what actions governments should take in helping citizens have access to computers may be answered by applying Rawls's principles of justice.<sup>2</sup> The application level is where moral deliberation takes place. Usually this involves the joint consideration of

---

<sup>2</sup> See for further discussion Van den Hoven (1997), who discusses models for applying moral theory in applied ethics.

moral theory, moral judgments or intuitions and background facts or theories, rather than a slavish application of preexisting moral rules.

Whereas computer ethics research at the theoretical level only requires philosophical expertise and may be carried out by philosophers, this is not so for research at the disclosure and application levels. Research at the disclosure level often requires considerable knowledge of the technological aspects of the system or practice that is studied, and often also requires expertise in social science for the analysis of the way in which the functioning of systems is dependent on human actions, rules and institutions. So ideally, research at the disclosure level is a cooperative venture between computer scientists, social scientists and philosophers. If not, it should at least be carried out by researchers with an adequate interdisciplinary background.

Research at the application level may be argued to be a philosopher's job again, as applying moral theory (e.g., weighing moral principles against considered moral judgments) seems to make an appeal to mostly philosophical skills (Van Den Hoven, 1997). However, even if bringing moral theory in agreement with moral judgments, empirical facts, scientific claims and other relevant sources of information is a activity that mostly appeals to philosophical skills, the information that must be processed in this task largely of a nonphilosophical kind. Philosophers engaged in this activity must therefore have a solid grasp of the social, legal and technical aspects of the technology or practice on which they are to pass moral judgments, or should opt to work with experts in these areas.

## **5. CONCLUSION**

Disclosive computer ethics constitutes a much needed approach in computer ethics that deviates from traditional approaches in applied ethics that usually focus on morally controversial practices and neglect embedded normativity in technological systems and practices, and still often concentrate on formulating and applying moral theory. As has been argued, disclosive computer ethics should preferably not be theory-driven, should be multi-level and interdisciplinary and should focus on four key values: justice, autonomy, democracy, and privacy. The proposed disclosive method may well be generalized to other areas of applied ethics in which technology plays an important role.

## References

- Agre, P. and Mailloux, C. (1997) Social Choice about Privacy: Intelligent Vehicle-Highway Systems in the United States, in *Human Values and the Design of Computer Technology* (ed. B. Friedman), Cambridge University Press, Cambridge.
- Akrich, M. (1992), 'The de-scription of technical objects,' In Bijker, W. & Law, J., eds., *Shaping Technology/Building society: Studies in Sociotechnical Change*, Cambridge, MA, MIT Press.
- Blanchette, J. (1998). 'On the Social Discourse of Cryptology.' Paper presented at CEPE'98, London School of Economics and Political Science, 14-15 December.
- Brey, P. (1997). 'New Media and the Quality of Life,' *Society for Philosophy and Technology Quarterly* 3:1, 1-23.
- Brey, P. (1998). The Politics of Computer Systems and the Ethics of Design, in *Computer Ethics: Philosophical Enquiry* (ed. J. van den Hoven), Rotterdam University Press, Rotterdam.
- Brey, P. (1999). 'Worker Autonomy and the Technological Drama of Computer Networks,' *Journal of Business Ethics* 22:1, 15-25.
- Dworkin, G. (1988), *The Theory and Practice of Autonomy*, Cambridge, Cambridge University Press.
- Feenberg, A. (1999). *Questioning Technology*. London and New York: Routledge.
- Friedman, B. and Nissenbaum, H. (1997) Bias in Computer Systems, in *Human Values and the Design of Computer Technology* (ed. B. Friedman), Cambridge University Press, Cambridge.

- Hill, T. (1991), *Autonomy and Self-Respect*. Cambridge, Cambridge University Press.
- Hoven, J. van den (1997). 'Computer Ethics and Moral Methodology.' *Metaphilosophy* 28: 3, 234-248.
- Introna, L. and Nissenbaum, H. (1998). 'Notes on the Politics of Search Engines.' Paper presented at *CEPE'98*, London School of Economics and Political Science, 14-15 December.
- Johnson, D. (1997). 'Is the Global Information Infrastructure a Democratic Technology?' *Computers & Society* 27: 20-26.
- Moor, J. (1985). 'What is Computer Ethics?' *Metaphilosophy*, 16, 266-275.
- Nissenbaum, H. (1998). 'Can We Protect Privacy in Public?' in *Computer Ethics: Philosophical Enquiry* (ed. J. van den Hoven), Rotterdam University Press, Rotterdam.
- Pfaffenberger, B. (1992), 'Technological dramas,' *Science, Technology, & Human Values* 17: 282-312.
- Rawls, J. (1971), *A Theory of Justice*, Cambridge, MA, Harvard University Press.
- Sclove, R. (1995), *Democracy and Technology*, Guilford Press, New York.
- Winner, L. (1980), 'Do Artifacts have Politics?' *Daedalus* 109: 121-136.