# Method in Computer Ethics: Towards a Multi-Level Interdisciplinary Approach

*Philip Brey*

This is a preprint version of the following article:

Brey, P. (2000). 'Method in Computer Ethics: Towards a Multi-Level Interdisciplinary Approach,' *Ethics and Information Technology* 2:3, 1-5.

## Abstract

This essay considers methodological aspects of computer ethics and argues for a multi-level interdisciplinary approach with a central role for what is called disclosive computer ethics. Disclosive computer ethics is concerned with the moral deciphering of embedded values and norms in computer systems, applications and practices. In the methodology for computer ethics research proposed in the essay, research takes place at three levels: the disclosure level, in which ideally philosophers, computer scientists and social scientists collaborate to disclose embedded normativity in computer systems and practices, the theoretical level, in which philosophers develop and modify moral theory, and the application level, that draws from research performed at the other two levels, and at which normative evaluations of computer systems and practices takes place.

## 1 Computer Ethics: Aim, Scope and Method

This essay considers the role of method in computer ethics and proposes that a particular methodology for doing computer ethics that is multi-level and interdisciplinary and assigns a special role to the moral deciphering of embedded

normativity or values in computer systems, application and practices. In this section, the methodological issues for computer ethics are introduced by relating them to the aim and scope of computer ethics. In section 2, a particular approach within computer ethics is discussed, which is called disclosive computer ethics. It is an approach that centers on the deciphering of embedded normativity in computer systems, application and practices. In section 3, the multi-level interdisciplinary methodology for doing computer ethics that is the topic of this paper is outlined, building on the discussions in the previous sections. The paper ends with a brief concluding section.

In his by now classical essay "What is computer ethics?" Jim Moor proposed that the central aim of computer ethics is to formulate policies to guide individual and collective action in the use of computer technology [Moor, 1985]. I agree with this proposal, with the addition that not just the use of computer technology, but also other practices that involve computing technology, such as its development or management, require the formulation of policy guidelines. When we conceive of computer ethics in this way, it is clear that it is a branch of *applied ethics*. Whereas its counterpart, theoretical ethics, is concerned with general aspects of morality, applied ethics is concerned with the study of morality in particular domains of human practice. Moreover, the aim of applied ethics is not merely to arrive at well-supported moral analysis, but also to use such analyses to affect the discourse, policies and practices that are prevalent in its domain of study.

The *scope* of computer ethics includes individual and collective practices that somehow essentially involve computers. This includes practices like the use, development,

regulation, management, advocacy and advertisement of computer technology. Also included should be the products of such actions, e.g., computer systems and software, manuals, advertisements, and laws and policies regulating the use of computers. These products deserve special mention because their moral properties may be analyzed independently from a consideration of the actions that have lead to them. For example, a law regulating the copying of software may be analyzed for its moral content independently of an analysis of any actions that led to the adoption of the law.

Having defined what I see as the aim and scope of computer ethics, I will now proceed to discuss various *method*s used within it, with the aim of identifying different research activities involved in contemporary computer ethics research. To begin with, computer ethics, like other branches of applied ethics, often involves the application of existing moral theory to practices that are under study [Van Den Hoven 1997]. So *the application of moral theory* is certainly one of the central activities in computer ethics. 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.

Applying moral theory is only part of what computer ethicists do, however. As Jim Moor [1985] has pointed out, the changing settings and practices that emerge with new computer technology may yield new values, as well al require the reconsideration of old values. There may also be new moral dilemma's 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.

Part of the work done in computer ethics is the development of ethical theory and its application to practices involving computer technology. Both these activities are *normative*, in that they are concerned with proposing, defending, analyzing or applying normative concepts and principles. I want to claim, however, that a large part of the research in computer ethics is not normative in this sense, but is instead *descriptive*: it is concerned with describing aspects of reality and with proposing, defending, analyzing or applying descriptive concepts and principles.

The importance of descriptive research has been noted to some extent by Jim Moor [1985], who has claimed that "much of the important work in computer ethics is devoted to proposing conceptual frameworks for understanding ethical problems involving computer technology." (p. 266). Moor clearly holds that a large part of the conceptual work needed for doing computer ethics is found in the analysis of descriptive concepts. For example, Moor holds that to arrive

at a policy for protecting computer programs, descriptive conceptual questions must first be answered such as "What is a computer program?" and "Can programs really be owned?". So a third important research activity for computer ethics is *conceptual analysis of descriptive concepts and adequate description of relevant empirical facts.*

## 2 Hidden Morality and Disclosive Computer Ethics

I want to argue that there is still a fourth important research activity in computer ethics, one that has not been recognized sufficiently by Moor. Moor seems to presume that computer ethics is in large part about solving preexisting moral problems. He claims: "A typical problem in computer ethics arises because there is a policy vacuum about how computer technology should be used." In such a case, the work that is to be done is the conceptual clarification and description of the practice that generates the moral problem. However, I want to claim that a large part of work in computer ethics is not about the clarification of practices that have already generated moral controversy, but rather about *revealing the moral import of practices that appear to be morally neutral*. Many designs and uses of computer systems, I want to claim, have important moral properties, that remain hidden because the technology and its relation to the context of use are too complex or are insufficiently well-known. It is part of the job of computer ethics to make computer technology and its uses transparent, in a way that reveals its morally relevant features.

The notion that computer technology can have moral properties is an extension of the notion that it can have political properties [Winner 1980; Sclove 1995; Pfaffenberger

1992; Akrich 1992]. 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. Many such structural effects of technology may be analyzed from a moral point of view.

Much recent work in computer ethics is centrally concerned with the moral deciphering of computer technology. Friedman & Nissenbaum [1997], for example, is a study of bias in computer systems (see also Brey [1998]). 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. 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 are [Nissenbaum, 1997], revealing the moral importance of practices of registering public information, [Blanchette 1998], revealing the importance of trust relations in cryptographic protocols and payment mechanisms, [Introna & Nissenbaum, 2000], who decipher the hidden politics of search engines, [Agre and Mailloux, 1997], who reveal the implications for privacy of Intelligent Vehicle-Highway Systems, and [Tavani, 1999], who analyzes the implications of data mining for privacy. The major contribution of all this research 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.

The importance of this mode of analysis in computer ethics justifies the introduction of a label by which it may be named. I propose to call this approach *disclosive computer ethics*. Disclosive studies in computer ethics are hence studies concerned with disclosing and evaluating embedded normativity in computer systems, applications and practices.

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 makes 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.

Finally, the question should be raised *what* moral values and norms are to be studied in disclosive analyses. I propose they should preferably be moral values and norms that are broadly supported throughout society, because computer systems or applications that violate them are clearly morally unacceptable. Some values that fit this criterion that have successfully been investigated in past studies in disclosive computer ethics are justice (or fairness), autonomy (or freedom), democracy, and privacy. So for example, a disclosive study of fairness in the design of an electronic credit rating system would start with a broad, common-sense definition of fairness and then proceed to investigate if there

are groups that are treated unfairly by the system according to the definition of fairness used.

## 3 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. This was identified in section 1 as one of the core tasks of computer ethics. This rather fundamental research, on issues like informational privacy or the relation between distributive justice and information, may be motivated by new practices involving computer technology, and may use concrete examples to support its claims, but aims to come to generalizations that abstract from specific technologies or practices. 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.

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 technical expertise, and often also require 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. Or else, it should 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.

The above three-layer model applies to disclosive approaches in computer ethics. Nondisclosive computer ethics normally follows a two-stage model that only includes the theoretical and application levels. Nondisclosive approaches are typically concerned with issues where it is already clear that the technologies or practices involved raise moral questions, and the aim is to try to answer these questions. In such studies, the technologies and practices are usually fairly transparent, but resolving the moral issues they raise turns out to be a challenge. For example, in studies on the ethics of anonymous speech on-line, it will usually be clear in advance what practices are at issue and what role technology plays in making them possible, so there is not

much work to be done at the disclosure level. Instead, most work will typically be done at the application level, in weighing and combining existing moral theory (e.g., on the ethics of anonymity) with the specifics of the case at hand.

## 4 Conclusion

Disclosive computer ethics constitutes a much needed approach in computer ethics that deviates from traditional approaches in applied ethics that usually 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 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.

The methodology required for disclosive computer ethics is a multi-level interdisciplinary one, in which research takes place at three levels: the disclosure level, in which philosophers, computer scientists and social scientists collaborate to disclose embedded normativity in computer systems and practices, the theoretical level, in which philosophers develop and modify moral theory, and the application level, at which individuals with good philosophical skills and a broad relevant background knowledge work on normative evaluations of computer systems and practices, drawing from research performed at the other two levels. Nondisclosive computer ethics research does not involve a disclosure level, but is a two-level process,

involving an application level and a theoretical level. It is not involved with revealing moral issues, but has its focus on an attempt to (further) clarify and resolve them.

## 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.

Akrich, M. (1992), 'The de-scription of technical objects,' In Bijker, W. & Law, J., eds., *Shaping Technology/Building society: Studies in Sociotechnical Change*, 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, *Techné: Journal of the Society for Philosophy and Technology*, 3, 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.

Brey, P. (1999), Worker Autonomy and the Drama of Digital Networks in Organizations, *Journal of Business Ethics*, 22, 15-25.

Dworkin, G. (1988), *The Theory and Practice of Autonomy*, Cambridge, Cambridge University Press.

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.

Hill, T. (1991), *Autonomy and Self-Respect,* Cambridge University Press.

Hoven, J. van den (1997), Computer Ethics and Moral Methodology, *Metaphilosophy*, 28, 234-248.

Introna, L. and Nissenbaum, H. (2000), The Politics of Search Engines, *Preferred Placement. Knowledge Politics on the Web* (ed. R. Rogers), Jan van Eyck Akademie Editions, Maastricht.

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.

Pfaffenberger, B. (1992), Technological dramas, *Science, Technology, & Human Values*, 17, 282-312.

Rawls, J. (1971), *A Theory of Justice*, Harvard University Press.

Sclove, R. (1995), *Democracy and Technology,* Guilford Press, New York.

Tavani, H. (1999), Informational Privacy, Data Mining, and the Internet, *Ethics and Information Technology*, 1, 137-145.

Winner, L. (1980), Do Artifacts have Politics? *Daedalus,* 109, 121-136.