This is a preprint version of the following article:
Brey, P. (1998). 'The Politics of Computer Systems and the Ethics of Design.' In *Computer Ethics: Philosophical Enquiry,* ed. M.J. van den Hoven , Rotterdam University Press, p. 64-75.

THE POLITICS OF COMPUTER SYSTEMS AND THE ETHICS OF DESIGN

Computer ethics has mostly been concerned with moral issues relating to the use, regulation, and social implications of new information technology. Few studies have been concerned with the design of computer systems, and those that are often have a narrow scope. In this essay, computer systems are analyzed as social structures, harboring design features that have political implications. The aim of this essay is to present a more comprehensive moral perspective on the design of computer systems that takes their political features into account. Analyses are presented of two classes of political features: user bias and user constraint (political properties in computer systems that affect users) and informational biases (political biases in the information function of computer systems). These features are evaluated in the context of a Rawlsian moral framework.

1 Introduction

In past and current research in computer ethics, two types of moral issues take center stage. The first includes issues concerning *individual morality*, or the morality of individual action. How, for example, should hacking be morally evaluated? How about software piracy or on-line conduct? What professional standards should computer professionals uphold, and what are their responsibilities and obligations to employers, clients, and society? Then there are issues that concern *social morality*, or the morality of collective action. These issues concern ways in which society as a whole should regard or respond to particular social implications of information technology. What is the moral importance of privacy, for example, and how should computer use be regulated to protect privacy? What is the moral importance of equal access to computer systems, and how should equal access be ensured? How do computers affect the quality of work conditions and personal life, and what should be done to minimize its negative impacts?[1]

What both types of issues have in common is that they usually somehow presuppose the existence of certain computer hardware and software, and then ask how this equipment ought to be *used*, or how its use should be *regulated*. What both tend to take for granted is the *design* of the computer systems in question. The few studies in computer ethics that do focus on design, moreover, often have a narrow focus. They focus on professional issues, like the obligations of designers to employers and clients, the morality of plagiarism, and liability for software errors, or focus on design standards relating to rather confined topics like health and safety, security, and system reliability.[2] They do not consider many of the broader social and political implications that computer systems designs may have.

The aim of this study is to offer a systematic perspective on computer systems design that addresses its moral relevance. This moral perspective goes beyond design features relating to health and safety, security, and system reliability, to include design features that have *political* import. These

[1] See, e.g., [Johnson, 1993], [Forester and Morrison, 1994], [Johnson and Nissenbaum, 1995], and [Spinello, 1995].
[2] See, e.g., the relevant chapters in the works cited in note 1.

are design features that contribute to social injustice, the formation of coercive power relations, and the curtailment of human freedom and autonomy.

At least since Langdon Winner's famous article 'Do Artifacts Have Politics?' [Winner 1980], it is recognized that technical artifacts may harbor political properties. Winner's best known example of artifacts with political properties tells of low-hanging overpasses on Long Island, that were built at a height that prevented the passage of buses. In this way, poor people and blacks from the City, the main users of public transit, were effectively kept away from Long Island and its public parks. As Winner explains, technological artifacts and systems function much like laws, by constraining behavior and serving as frameworks for public order.

Richard Sclove makes the same point by identifying technical artifacts as social structures [Sclove, 1995]. Sclove defines the overall 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.' (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 social structures. 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.[3]

If technologies function like social structures, then they should be morally evaluated just like other social structures. Moral theories of social structure are found in the work of social ethicists like Rawls, Dworkin, and Waltzer. Although these theorists do not usually identify technologies as social structures, their work can usually be readily extended to accommodate technology. There is hence no need, I claim, to construct radically new theories of social ethics to accomodate information technology (although I will argue later on that the application of existing moral theory to computer systems design does require a special form of descriptive analysis). In the next section, I will propose a general social ethic based on Rawls' theory of justice. In subsequent sections, this social ethic will be applied in the analysis of political properties of computer systems, viewed as social structures.

2 Social structure and primary goods

The two social values most often extolled in democratic nation states are probably those of *individual autonomy* (or *freedom*), and *justice*. The notion of 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.

John Rawls calls goods that are prerequisite to carrying out one's life plan *primary goods* [Rawls, 1971]. According to Rawls, primary goods are things that every rational human being is presumed to want because they have a use whatever a person's plan of life. Rawls distinguishes between *natural* primary goods, goods that people possess largely independently of the way in which society is structured, and *social* primary goods, the distribution of which is strongly dependent on the basic structure of society. Natural primary goods include health and vigor, intelligence and

[3] Extended discussion and examples of these properties of technologies are found in [Winner, 1980], [Sclove, 1995], [Pfaffenberger, 1992] and [Akrich, 1992].

imagination. Social primary goods include rights and liberties, powers and opportunities, income and wealth, as well as the social bases of self-respect.

Rawls employs the notion of a primary good in devising his theory of justice. 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 goods. In particular, society should not promote the unfair distribution of primary goods, as these are essential for individuals to carrying out their life plans. Rawls has formulated this insight in his famous two principles of justice:

> Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all."
> and
> Social and economic inequalities are to be arranged so that they are both:
> (a) to the greatest benefit of the least advantaged [...], and
> (b) attached to offices and positions open to all under conditions of fair equality of opportunity. [Rawls, 1971, 302]

These two principles, henceforth named together 'the principle of justice,' are an elaboration of the basic insight that the distribution of primary goods in a society should be fair.

I will adopt the principle of justice, but will also adopt another moral principle based on Rawls' notion of a primary good, which I will call the principle of autonomy:

> Social structures should be arranged in a way that optimizes, given present knowledge, the total system of primary goods available to all.

The principle of autonomy is meant to say that it is morally imperative for those individuals and organizations that are implicated in the arrangement of social structures to ensure not only that these structures are just, but also that they promote overall human autonomy, and do not unnecessarily constrain the quantity of primary goods available to individuals. The principle would, for example, prescribe that providers of drinking water should undertake reasonable efforts to purify their water, even if unpurified drinking water does not violate the principle of justice.

How do these two principles apply to computer systems, viewed as social structures? First, computer systems are elements of the information infrastructure of a society. Their intended function is an *information function*, consisting in the storage, transmission, or processing of information. Van den Hoven [1995] has argued that information should be counted as a primary good, because it often has an important instrumental role for individuals in carrying out their life plans. Because of the increasing importance of computer systems as providers of information, access to computer systems is therefore quickly adopting the status of a primary good. It is morally imperative, then, that both computer systems and their surrounding social structures are arranged so as to facilitate equal access to the sundry information functions of computer systems.[4]

Many of the implications of computers for the distribution of primary goods in a society are however latent and unintended. Computers may, for example, lead to unemployment and deskilling, introduce hierarchical power relations, or harm privacy. Computer systems and surrounding social structures should therefore not just be designed to ensure equal access to prospective users, they should also be structured to minimize negative social impacts that are latent and unintended.

---

[4] Note that individuals do not only use computers to receive information, but also use them to communicate information to others. The ability to communicate information to others is also a good candidate for a primary good, because communication is, like information, often instrumental for individuals in carrying out their life plans.

3 Political ergonomics of computer systems

The most important ingredient for a moral assessment of the design of computer systems is not, I claim, a moral theory of the sort outlined above. It is, instead, a descriptive analysis of the way in which designs come to function as social structures with political implications. Langdon Winner has even proposed a separate field of study for this purpose, *political ergonomics*, which should have as its aim 'to decipher the design features - both general and specific, large and small - of technological devices for their social and political significance.' [Winner 1995, 162]. Equipped with this knowledge, it should be possible 'to anticipate and guide the contribution that a particular device or system makes to the quality of political society' [ibid.].

Political ergonomics would therefore be a prerequisite to full-blown ethical analyses of particular technological designs. It would generate analyses of the fit between the design features of technological devices and the society that they are part of, as well as constructive tools for the transformation of existing sociotechnical orders. Moral theory could have a subsequent role in the evaluation of existing designs and in the guidance of future technological development. In practice, however, this separation between political ergonomics and ethical analysis may not be entirely feasible. To analyze design features for their social and political implications, one must first determine what social and political implications are relevant, and this already implies a moral position. In my own analysis of the design features of computer systems, I will take the two moral principles of section two as explicit points of departure: I will analyze ways in which the design features of computer systems affect the overall distribution of primary goods in society.

My analysis of the design features of computer systems is based on several theoretical assumptions. One of these is that computer systems can be understood as social structures. Another is that, as social structures, computer systems have political properties. However, political properties of artifacts always exist *relative to* a surrounding social structure in which these artifacts are integrated. For example, the overpasses in Long Island would not have their political properties if they were not part of a larger social structure that included particular systems of transportation and particular distributions of wealth and income. If a poor fit exists between a technology and a surrounding social structure, one choice is to redesign the technology, but an alternative choice is to reengineer the surrounding social structure (e.g., by redistributing wealth and income, or by building lower buses).

It is moreover assumed that social structures are malleable, and adapt to technological change in ways that are not always predictable. Newly introduced technology brings along opportunities and threats, to which people respond in ways that may change the surrounding social structure, if not the technology itself, and thereby affect its impacts (cf. [Pfaffenberger, 1992]). As a result, the political properties of technologies are not fixed but depend on social responses. Especially difficult to predict are macro-level effects that depend on the behavior of many different actors. In these cases, extensive knowledge would be required of the surrounding social structure, and complex models would have to be constructed of expected responses of different actors to newly introduced technology. Political ergonomics then faces all of the usual methodological problems of technology assessment.

Initial studies in political ergonomics should therefore perhaps focus on political properties of artifacts that do not depend on the actions of many actors. I propose that two classes of political properties are relatively easy to study. They are design features that impact users (*user biases* and *user constraints*), and political properties implicated in the intended function of the artifact (*functional biases*). The assessment of user bias is often straightforward because there is only one major independent variable: the user. The assessment of functional bias is often straightforward as well, because functional biases in an artifact usually manifest themselves whenever it is used (roughly) according to its intended function. For example, a flight scheduling program that gives priority to flights with low flight numbers will display this bias whenever the system is used. In this essay, therefore, the focus will be on user bias, user constraint, and *informational bias*, which is functional bias in the information function of computer systems.

Of the few studies that qualify as studies in the political ergonomics of computer systems according to Winner's criteria, most seem to be concerned with improving the overall fit between a computer and its immediate social environment (e.g., [Norman, 1990, 1993], [Winograd and Flores, 1986]), a concern that conforms to the principle of autonomy. I am aware of only one study that systematically considers ways in which computer systems violate the principle of justice. This is a study of 'bias in computer systems= by Friedman and Nissenbaum [1994].

Friedman and Nissenbaum distinguish three basic types of bias in computer systems: *preexisting bias*, which arises from the values and attitudes that existed prior to the design of the system, *technical bias*, which arises from technical constraints or considerations, and *emergent bias*, which arises when the social context in which the system is used is not the social context intended by its designers. Their typology is revealing and well supported, and is especially helpful in locating different phases in design trajectories in which biases are infused. However, their typology classifies biases by their *causes*, whereas I will be concerned with constructing a typology of biases on the basis of their political *effects*. My typology of biases in computer systems will therefore crosscut Friedman and Nissenbaum's typology.

4 User bias and user constraint

User biases are design features of computer systems that disadvantage some of their potential users. Design features that are against the interests of all users are called *user constraints*. They will now be discussed in turn.

4.1 User bias

User biases are features of computer systems that pose a disadvantage to particular users or user categories. A user bias may be called *serious* if it has as a consequence that the net gain (loss) in primary goods through use of the system by targeted users is significantly smaller (larger) than that of other users. When a user bias is serious, and no alternative systems are available to targeted users without a significant sacrifice in primary goods, the system violates the principle of justice, and the system may be called unjust. A moral imperative then exists to redesign the system, or, alternatively, to redesign its surrounding social structure to eliminate the injustice.

Three major types of user bias can be distinguished: user exclusion, selective burdening of users, and user-directed informational bias.

(1) User exclusion. This is the selective exclusion of users from using some or all of the functional features of a computer system, because they lack required qualities or competencies.

(2) Selective burdening of users. This is the selective presence of disadvantages that some users have when using a computer system, because they lack certain qualities or competencies. The lack of these qualities or competencies makes using the system more taxing in terms of time, effort, or resources, or introduces elements of personal risks.

User exclusion and selective burdening of users are both biases that disadvantage prospective users that in some way do not fit the profile of an 'ideal' or 'normal' user. In user exclusion, this results in the system being wholly or partially unusable, whereas in selective burdening, the system can be used, but only at a price.

There are three principal ways in which prospective users of computer systems may deviate from an 'ideal' or 'regular' user. First, prospective users may have a physical or mental *disability* that makes the use of a system impossible or at least burdensome. The most common disabilities that make most computer systems difficult or impossible to use are blindness and partial sightedness, manual disabilities, memory impairments, and dyslexia. Some of these, such as blindness and manual disabilities, rule out the use of most computer systems, whereas others, such as dyslexia and nearsightedness, impose heavy burdens on users.

Now that computer systems increasingly qualify as carriers of primary goods, it is increasingly important for people with disabilities to have access to them, in a way that does not impose unacceptable burdens. Ideally, the design of mainstream computer systems would accomodate as many people with disabilities as possible. When this is not feasible, alternative computer systems should exist, or computer systems should be compatible with the use of peripheral devices, designed for people with disabilities, to facilitate access.

The U. S. Telecommunications Act of 1996 provides an example of how these moral imperatives can be put into practice. Section 255, which regulates access to telecommunications systems by person with disabilities, states that manufacturers of telecommunications equipment have to ensure that equipment is designed, developed, and fabricated to be usable by individuals with disabilities, if this is readily achievable. If this is not readily achievable, manufacturers have to ensure that the equipment is compatible with existing peripheral devices commonly used by individuals with disabilities, again only if this is readily achievable.

Besides requiring able-bodied users, most computer systems also require that users possess *computer skills*. This is a second type of attribute that individuals may lack. Relevant computer skills normally include skills for operating keyboards and reading displays, and a practical understanding of the operating system and particular software programs. In the public debate surrounding the current information revolution, the worry has been voiced that there is an emerging social divide between computer literates (individuals who are adequately equipped with computer skills) and computer illiterates (individuals who lack such skills). The implicit assumption in the expression 'computer literacy' is that computer skills are becoming as important as reading and writing skills in contemporary society. In many societies, linguistic illiterates are second-rate citizens, who are denied many of the powers and opportunities of literate citizens. The worry is that a new class of second-rate citizens is emerging, consisting of computer illiterates.

Again, given the growing importance of access to computer systems, there is a moral imperative for a society to ensure that all of its citizens are given the opportunity to have such access. If it turns out that a significant segment of the population lacks required computer skills, then there are two obvious ways to remedy this situation. The first is to simplify the design of computer systems, so that fewer computer skills are needed to use them. A case can be made that societies nowadays have a moral obligation ensure that computer systems are designed to be easy to use, and accomodating to individuals with different >cognitive styles=. A second strategy is to improve the computer skills of citizens, through education and training. Individuals must be given the opportunity to acquire computer skills, and the educational system must be structured to accomodate all who enter through it in this respect.

A final category of 'ideal' attributes that prospective users may lack are *external resources*, usually including relevant computer hardware and software, that are necessary to use some computerized function or service. Individuals may be barred from or burdened in accessing or using certain computerized services and functions, because they are not equipped with the right software or hardware. For example, many sites on the World Wide Web favor access by fast computer systems with up-to-date software, as do sites that make heavy use of graphics that burden slower computer systems with long download times.

Because computerized information services and functions increasingly have the status of a primary good, it is imperative that they are easily available to citizens, and that high-quality access to them is not denied because they lack the appropriate equipment. Nowadays, lower-income households often cannot afford adequate computer equipment, but even middle-income households often find that their computer equipment ages too quickly to maintain high-quality access to information services, or is just not compatible with certain software or hardware. These problems can be partially solved by measures external to the technology, such as partial financing of computer equipment for lower-income households by governments, the funding of public access locations, and by ensuring that noncomputerized equivalents to those information services that are most vital to citizens are available to them. Part of the solution must come from the technology itself, however. Computer hardware and software and on-line services ought to be designed to ensure that important computerized information services are easily accessible to the vast majority of computer owners.

(3) User-directed informational bias. Computer systems with a user-directed informational bias have functional features that turn out to be less compatible with the interests and values of some users. User-directed informational bias differs from selective burdening of users in that targeted users need not find operating the system more taxing, but instead find that the system's information function is less useful or less agreeable to them. Examples of user-directed informational bias are found in instances of educational software that have more appeal to boys than girls, or that makes use of competitive elements, but is used by students with cultural backgrounds that largely eschew competition and promote cooperative endeavors.[5] Another example would be an expert system that helps investors build portfolio's, that works well for wealthy investors but gives poor advice when smaller amounts of money are entered.

User-directed informational biases violate the principle of justice when the information function provided by the system is an important one and it is burdensome or impossible for prospective users to switch to an alternative system more compatible with their values or interests. Computer systems should be designed so that they accomodate for the interests and values of prospective users when reasonable alternatives are not available to them. When biases cannot reasonably be eliminated, prospective users should be made aware of them.

4.2 User constraint

User constraints are attributes of computer systems that work to the disadvantage of all users, by imposing unnecessarily penalties on their primary goods. The more obvious biases of this kind are ones that affect the health and safety of users. However, computer systems frequently also harbor other, less obvious biases that have implications for the autonomy of users.

(1) Presence of health and safety hazards. These are implications for physical and mental health and personal safety that all users of computer systems are subjected to. Now that many people spend a significant proportion of their workweek and leisure time behind computer screens, preventive measures to limit such hazards have gained in importance. Relevant hazards include repetitive strain injury, which is a painful and disabling affliction to the muscles caused by excessive use of keyboards, eyesight problems because of radiation emitted by video displays, as well as other symptoms such as fatigue, dizziness, and insomnia. Obviously, it is a moral imperative for designers to attempt to design more ergonomic computer systems in which such hazards are minimized.

(2) Presence of monitoring functions. Computer systems are sometimes designed to enable or allow for the monitoring of the activities or the files of the user by other individuals, such as managers, system operators, and coworkers. A word processor used by an employee may, for example, be designed to allow management to measure the number of keystrokes performed on a working day, or to access stored files. There will be many instances in which the negative impacts such monitoring has on

[5] Both these examples are taken from [Friedman and Nissenbaum, 1994].

users are negligible. However, as documented in [Forester and Morrison, 1994], office workers who are subjected to computerized monitoring are more likely to suffer from stress and ill health, and suffer from declining morale.[6] Computerized monitoring undermines privacy and reduces the work autonomy of employees by creating a power relation between the user and the monitoring agent.[7] Computer systems that include monitoring functions are therefore morally controversial.

(3) Dependency bias. A computer system exhibits dependency bias when it is structured to make its users dependent on others who serve as mediators between them and the information functions provided by the system. A computer system may for instance be structured to assign a powerful role to the system operator, making access to system functions dependent on his or her authorization, and giving him or her the power to influence or interrupt user operations. A system may also be so difficult to operate, and so full of glitches, that its operation requires constant intervention by, and advice from, computer professionals. The autonomy of users is consequently eroded, and users may come to feel dependent and constrained in their actions.[8]

(4) Machine-centeredness. A well-designed computer system is experienced by its users as a natural extension of themselves that allows them to creatively put their ideas into action. A poorly designed system, in contrast, is experienced as an autonomous agent that is inflexible and imposes its own logic on the user. Users have only limited control over its operation, and must wrestle with the system to get it to do things that are helpful to them. Donald Norman calls computer systems in the first category human-centered, and systems in the second class machine-centered [Norman, 1993]. Norman argues that many technological artifacts, especially computer systems, suffer from a machine-centered approach to design, that puts the 'needs' and idiosyncrasies of machines before those of human beings, requiring users to conform to machines, rather than machines to their users.

Computer systems are often designed around a machine logic that requires precision, orderliness, absence of error, and instructions that are meaningful for the computer but not for the user, and has little tolerance for the fact that users do make errors, and are often disorganized and forgetful. The principle of autonomy favors a human-centered approach over the still prevailing machine-centered approach to the design of computer systems. In this approach, the needs and idiosyncrasies of the user, and not those of the system, are central. Systems are to offer choice rather than constrain it, allow for skillful work, and keep users in control of important operations of the system.

5 Informational bias

Computer systems are information-processing systems: their function is to store, retrieve, transmit, manipulate, or generate information. Functions of this sort were called information functions. An *informational bias* is a bias in one or more of the information functions of a computer system. It is a design feature because of which a system unfairly disadvantages some of its *informational stakeholders*, which are individuals, groups or organizations that in some way have a stake in the way in which the information function of a computer system is designed. One type of informational bias has already been discussed: user-directed informational bias, a bias that targets users. The emphasis in this section will be on informational biases that target stakeholders other than the user.

Informational bias is a type of bias found in any device that embodies information functions. It is not just found in computer systems, but also in artifacts like books, notebooks, planners, filing

[6] Forester and Morrison, 1994, p. 211-213.
[7] [Barker and Downing, 1985] describes how computerized monitoring has reduced the work autonomy of typists. See also [Hawk, 1994].
[8] Processes of dependency creation through technology are discussed at length in [Akrich, 1992] and [Illich, 1973].

systems, telephones, televisions, and measuring instruments. The yellow pages, for example, contain (modest) informational biases by using larger print for the names and numbers of businesses that pay for this service, and by using alphabetized orderings that favor businesses whose names start with letters found early in the alphabet. An electronic yellow pages could harbor similar biases, but it may also contain informational biases unique to computer systems, such as biases in the structure of search algorithms.

There are two ways in which devices with information functions may come to exhibit informational bias. Informational bias is sometimes the result of biases in the *informational content* of the information that is stored in, or transmitted by, a device with an information function. Televised information may, for example, contain statements that go against the interests of some individuals or groups. However, informational bias often also results from the particular manner in which information is encoded, stored, transmitted, retrieved, and displayed: the *organizational* aspects of an information device. Biases that result from such organizational aspects may be called *organizational informational biases*. The importance of organizational aspects is highlighted in Marshall McLuhan's famous diction 'the medium is the message.' Many of the most serious biases in information devices are not found in the informational content of their data, but in the way in which these data are organized by the system.

For two classes of computer systems, database systems and knowledge-based systems, a consideration of organizational bias is especially important, because they often have identifiable informational stakeholders that have a significant stake in their organizational aspects. *Database systems* are computerized systems that are used to organize, store, and retrieve data. Information is normally retrieved by search mechanisms that match properties of stored data with properties of search data provided by users. Some search mechanisms look for simple matches between symbol strings, whereas others contain complex search algorithms. Informational stakeholders of database systems often include individuals or organizations that are in some way *represented* by data records contained in the database.

Knowledge-based systems are expert systems and other 'intelligent' computer programs that use a knowledge base consisting of rules to draw inferences from a body of data. Examples are medical expert systems that make diagnoses based on medical data, and legal systems that recommend a penalty based on the profile of a convicted offender.[9] The informational stakeholders of knowledge-based systems often include users, because users often have a personal stake in the information provided by the system. However, users may also have little or no stake in this information. For example, a bank employee who uses loan approval software to decide whether a client can receive a loan has no personal stake in the absence of bias in the software.

There are at least four organizational features of database systems (and knowledge-based systems) that can embody informational biases. First, the *selection of data* to be included in a system may be unfair in excluding data that are relevant to the intended information function of the system. For example, a program may present itself as an electronic yellow pages, but only include businesses who have paid for inclusion in the database. Second, biases may be contained in the *system of categories* used for categorizing data and for presenting options to the users. For example, a database system of a record retailer on the Internet may categorize records in categories like 'rock' and 'jazz,' but fail to include a 'rhythm & blues' category, instead distributing r&b records haphazardly over other categories, thereby disadvantaging both users who are r&b fans and record companies who issue r&b records.

Third, biases may be contained in *search and matching algorithms* and *inferential rules*. For

[9] Notice that knowledge-based systems may include databases containing ordinary data, and database systems may include knowledge-based search and storage mechanisms. The two types of systems are hence not mutually exclusive, and the boundaries between the two are fuzzy.

example, an automated credit advisor may exhibit bias by weighing an ethnic surname as in evaluating the creditworthiness of an applicant. A scheduling algorithm that schedules airplanes for takeoff may give priority to airlines with names low in the alphabet.[10] An electronic dating system may fail to match clients of different races, although they were not asked for their racial preferences. Less controversially, a search program for the World Wide Web may give high marks to Web pages that contain multiple occurrences of search terms, while failing to give weight to the occurrence of the search term in the name of the page.

Fourth, biases may be contained in the *display* of information. Database searches often result in a series of matches that are displayed as options for further action (e.g., airline flights, books, hyperlinks to Web pages, etc.) The order in which otherwise equal options are presented often works to the advantage of those options that are presented at the top of the list. This effect is especially strong when the display of a set of options requires multiple screens, in which case options that appear on later screens are often at a great disadvantage. The use of large print, color, flashing print, the provision of extra information on some option, the setting of default values, and the selectively provision of hyperlinks all introduce biases by singling out options for special attention.

Many informational biases found in computer systems are too mild to violate the principle of justice. However, information biases sometimes yield great injustices, especially when the informational function of the system is used to make decisions that have major effects on the lives of some individuals. It is in these cases imperative, and in other cases desirable, that informational bias is avoided in design, to the extent that this is possible.

6 Redistributing responsibility for computer systems design

Taking the problem of bias in computer systems more seriously would imply changes in current design practice. It might be thought that the above ethical analyses of bias in computer systems translate straightforwardly into a set of moral imperatives for their designers. Yet, this is not the case, for two reasons. First, designs are rarely the product of individual designers, but are usually the result of the activities of the members of a *design constituency* consisting of many individuals and organizations, such as design engineers, clients, managers, and regulators [Staudenmaier, 1989]. The responsibility for a technological design is the collective responsibility of a design constituency. The translation of collective responsibility to responsibilities at the individual level is a difficult task.

Second, it was remarked in section 3 that if a poor fit exists between a technology and a surrounding social structure, a better fit may be attained either by redesigning the technology, or by reengineering the surrounding social structure. Whether redesign of computer systems is morally imperative therefore depends on a moral comparison of the options of systems redesign and social reengineering. For these two reasons, then, a distribution of individual and collective responsibility cannot be derived directly from an ethical analysis of bias in technology. Such a distribution should, I suggest, be based on the outcome of a democratic process of social negotiation, that should yield a distribution of responsibilities that is fair to all parties involved, and that satisfies criteria of practicability, efficiency, and effectiveness.

[10] Both examples are taken from [Friedman and Nissenbaum, 1994].

REFERENCES

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.

Barker, J. and Downing, C. (1985), 'Word processing and the transformation of patriarchal relations of control in the office,= In MacKenzie, D. & Wajcman, J., eds., *The Social Shaping of Technology: How the Refrigerator got its Hum*, Philadelphia, Open University Press.

Forester, T. and Morrison, P. (1994), *Computer Ethics: Cautionary Tales and Ethical Dilemma's in Computing*, Cambridge, MA, MIT Press.

Friedman, B. and Nissenbaum, H. (1994), *Bias in Computer Systems*, Stanford, CA, CLSI Report No. CLSI-94-188. Also published in *ACM transactions on information systems*, 14, 1996, p. 330ff.

Hawk, S. (1994), >The Effects of Computerized Performance Monitoring: An Ethical Perspective,= *Journal of Business Ethics*, 13: 949-957.

Hoven, M. van den (1995), *Information Technology and Moral Philosophy*, dissertation Erasmus University, Rotterdam, The Netherlands.

Illich, I. (1973), *Tools for Conviviality,* Harper and Row, New York.

Johnson, D. (1993), *Computer Ethics*, 2nd ed., Englewood Cliffs, NJ, Prentice-Hall.

Johnson, D. and Nissenbaum, H. (eds.) (1995), *Computers, Ethics and Social Values*, Englewood Cliffs, NJ, Prentice-Hall.

Norman, D. (1990), *The Design of Everyday Things*, New York, Currency/Doubleday.

Norman, D. (1993), *Things that make us smart: Defending human attributes in the age of the machine*, Reading, MA, Addison-Wesley.

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.

Spinello, R. (1995), *Ethical Aspects of Information Technology*, Englewood Cliffs, NJ, Prentice-Hall.

Staudenmaier, J. (1989), > The politics of successful technologies,= in Cutliffe, S. & Post, R., eds*., In Context: History and the History of Technology*, Betlehem, PA, Lehigh University Press.

Winner, L. (1980), >Do Artifacts have politics?= *Daedalus* 109: 121-136.

Winner, L. (1995), 'Political ergonomics,' In Buchanan, R. & Margolin, V., eds., *Discovering Design: Explorations in Design Studies,* Chicago, University of Chicago Press.

Winograd, T. and Flores, F. (1986), *Understanding Computers and Cognition: A New Foundation for Design*, Norwood, NJ, Ablex.