Value-Sensitive Design:
A Research Agenda for Information Technology

A Report on the May 20-21, 1999
Value-Sensitive Design Workshop

August 23, 1999

Batya Friedman
School of Library and Information Science
University of Washington
batya@u.washington.edu

Formerly at:
Department of Computer Science
Colby College

In recent years, there has been a growing national awareness that information technology impacts our social lives not only positively but sometimes negatively. Correspondingly, impending initiatives have called on the government to conduct research to identify, understand, anticipate, and address the social and value-laden problems that arise from the fast-paced development and deployment of information technology (e.g., see the final report of PITAC, 1999). Value-Sensitive Design offers a promising answer to this call. Accordingly, this workshop was convened to frame a research agenda on Value-Sensitive Design.

Early interest in technology, values, and design emerged in the work of Mumford (1934), Wiener (1954), and Kling (1980). More recently, such work has been pursued by Friedman et al. (1996a, 1996b, 1997a, 1997b), Gotterbarn et al. (1997), Grudin (1999), Nissenbaum (1996, 1998a, 1998b), Shneiderman (1991), Suchman (1994), Winograd (1994), and others (see, e.g., Friedman's, 1997a, edited volume titled *Human Values and the Design of Computer Technology*). Following Friedman (1996a, 1997a), we refer to this emerging body of work as *Value-Sensitive Design*.

However, to date, the work on Value-Sensitive Design has been conducted with small budgets or in-between the cracks of other large projects with other agendas. If we are to make progress in developing the theory, methods, and practice of Value-Sensitive Design, then adequate resources must be devoted to this endeavor.

The workshop was funded by The Information and Intelligent Systems Division of the National Science Foundation and held at the National Science Foundation in Arlington, VA on May 19-20, 1999. The stated objective of the workshop was to clarify the definition, content, and scope for Value-Sensitive Design, and to enumerate a near-term (within the next five years) research agenda and corresponding priorities. In the context of addressing these issues, the workshop participants also identified relevant knowledge communities and exemplar projects, and clarified the unique identity of Value-Sensitive Design.

The workshop was organized by Batya Friedman of the University of Washington. Workshop participants included: Edward Felten of Princeton University, Batya Friedman of the University of Washington, Jonathan Grudin of Microsoft, Helen Nissenbaum of Princeton University, and Terry Winograd of Stanford University. As a group, the participants represented perspectives from cognitive science, computer ethics, computer security, computer-supported cooperative work, design, human-computer interaction, interaction design, philosophy, social psychological aspects of information systems, software development, and technology. The workshop was sponsored by: Darleen Fisher, Program Officer of the Special Projects and Networking Research Program; Rachelle Hollander, Program Officer of the Societal Dimensions of Engineering, Science, & Technology Program; Suzanne Iacono, Program Officer of the Computation and Social Systems Program; and Michael Lesk, Director of Computer & Information Science & Engineering. Darleen Fisher, Rachelle Hollander, and Suzanne Iacono attended the workshop and actively participated in the discussions. Michael Lesk was present for a brief presentation and discussion. Suzanne Iacono played a critical role in conceptualizing the workshop, and bringing it to fruition.

Section I: Clarifying the Definition, Content, and Scope of Value-Sensitive Design

Working Definition of Value-Sensitive Design

Value-Sensitive Design refers to an approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process.

Value-Sensitive Design is primarily concerned with values that center on human well being, human dignity, justice, welfare, and human rights. Specific values include trust, accountability, freedom from bias, access, autonomy, privacy, and consent. Value-Sensitive Design connects the people who design systems and interfaces with the people who think about and understand the values of the stakeholders who are affected by the systems. Thus, the theory and methods of Value-Sensitive Design are to be used in consort with other existing technical methods. Ultimately, Value-Sensitive Design requires that we broaden the goals and criteria for judging the quality of technological systems to include those that advance human values.

For purposes of this workshop, participants focused discussion on information technology.

What is unique about Value-Sensitive Design?

From a variety of perspectives, there has been increasing interest in and concern with the social implications of information technology. Efforts along these lines include social analyses of computing (e.g., Kling, 1980; 1996; Sproull & Kiesler, 1991; Suchman, 1987; Zuboff, 1988), computer ethics (e.g., Ermann, Williams, & Gutierrez, 1990; Forester, 1989; Johnson, 1994; Johnson & Nissenbaum, 1995; Ladd, 1989; Moor, 1985), participatory design (e.g., Bodker, 1991; Ehn, 1988; Greenbaum & Kyng, 1991; Kuhn, 1996; Schular & Namioka, 1993), and computer-supported cooperative work (Baecker, 1992; Blomberg, Suchman, & Trigg, 1996; Clement & Van den Besselaar, 1994; Grief, 1988). What, then, distinguishes Value-Sensitive Design from other socially-oriented approaches to information systems?

While recognizing valuable contributions from these other perspectives, participants in the workshop identified a constellation of features that is unique to Value-Sensitive Design:

- *Proactively Oriented toward Influencing Design*: While the theory and methods of Value-Sensitive Design can be applied to retrofitting problematic existing systems, Value-Sensitive Design is oriented toward influencing the design of information technology early in and throughout the design process. There is much to be learned in the service of developing Value-Sensitive Design from studying why systems fail, with the overarching goal to improve the quality of our information systems before these systems are deployed in society.
- *Carrying Critical Analyses of Human Values into the Design and Engineering Process*. Value-Sensitive Design is committed to design and engineering methodologies that bring critical analyses of human values into the design process.

- *Enlarging the Scope of Human Values.* Value-Sensitive Design shares a commitment with (a) Participatory Design to values of participation and democracy, and (b) Computer-Supported Cooperative Work to values of collaboration in the workplace. However, Value-Sensitive Design embraces a broader spectrum of human values that arise in human activity (e.g., in education, the home, commerce, online communities, and public life). As noted in the working definition, Value-Sensitive Design is primarily concerned with values that center on human well being, human dignity, justice, welfare, and human rights. Specific values include trust, accountability, freedom from bias, access, autonomy, privacy, and consent.
- *Broadening and Deepening the Methodological Approaches*. The emergent methods of Value-Sensitive Design draw on anthropology, organizational studies, design, psychology, philosophy, sociology, human-computer interaction, and software engineering. Specialized criteria and corresponding metrics (e.g., for informed consent online, deserved trust online) guide and measure the success of particular designs.

Knowledge Communities

Researchers from the following knowledge communities are likely to contribute to the development of Value-Sensitive Design:

- Designers
- Humanists (e.g., historians, philosophers)
- Social scientists (e.g., cognitive, developmental, and social psychologists; communications specialists; cultural anthropologists; economists; political scientists; and sociologists)
- Technologists (e.g., computer scientists and engineers, information scientists, new media specialists)

The workshop participants also identified a preliminary list of researchers whose work, broadly conceived, fits within Value-Sensitive Design. A list of these researchers can be found in Appendix 1.

Exemplar Projects

To provide a flavor for work in Value-Sensitive Design, participants identified a number of *exemplar* projects. These projects illustrate that Value-Sensitive Design can be non-intrusive in the design process and cost competitive in the marketplace. Each project is described briefly.

- Ben Shneiderman's (Shneiderman & Rose, 1997) redesign of an information system for Maryland's Department of Juvenile Justice led to his Social Impact Statement instrument. This instrument provides a systematic means to identify stakeholders, value implications, and effects from a proposed design early in the design process. The resulting Social Impact Statement can function as a statement of intended effects against which the success of the resulting system can be judged.
- John Thomas' (1997) work on ISO 9001 registration at NYNEX Science and Technology led to the systematic consideration and documentation of design practices to minimize bias and enhance access to communications technology. For

example, in the ISO 9001 registration document, Element 4.10 Inspection and Testing includes procedures to help ensure that a wide representation of possible users make up any test group and Element 4.4 Design Control includes considerations for people with special needs.

- Edward Felten's, Batya Friedman's, and Helen Nissenbaum's (in progress) work on human values and network browser security provides a "proof-of-concept" project for value-sensitive design methodology in which critical value analyses and user studies drive the technical design of web browser technology.
- Vicki O'Day's (Nardi & O'Day, 1999) work with *Pueblo* (an online school community) articulated value considerations – of trust, conflict resolution, and informed consent – appropriate to an online community for elementary school children. These value considerations were then used to influence the design of technical mechanisms to support the online community's social interactions.
- Jonathan Grudin's analysis of online calendar systems in a large organization led to the finding that increased visibility of personal information (i.e., daily commitments) can enhance a climate of trust.
- Phil Agre's (Agre & Harbs, 1994) analysis of privacy and intelligent-vehicle highway systems investigated how the existing technical strengths of information technology organizations influence proposed solutions to social problems, how (if at all) human values are accounted for in those solutions, and how those values are manifest in the resulting technical designs.
- John Tang's (1997) analysis of eliminating a hardware on/off switch on the microphone of a workstation points toward economic incentives for value-sensitive design and the need to better understand and articulate the economic consequences.
- Batya Friedman's and Helen Nissenbaum's (1996) retrospective analysis of bias in computer systems identified three categories of bias: pre-existing, technical, and emergent. In turn, each bias category can be applied to existing systems and systems under development to help identify and minimize potential bias.
- Deborah Johnson's and John Mulvey's (1993) work examined how values are incorporated into large-scale computerized models that are used for policy making purposes (e.g., by the Forest Service to make decisions about forest use based on computer models).

In addition, emerging technologies that come under the purview of Value-Sensitive Design include:

- Open-source process that supports democratization and participation by technologists. Large-scale open source projects work well in part because underlying technical mechanisms support coordination and communication among technical practitioners working on the project.
- Technical mechanisms designed to support anonymity online have been effectively used in email, ecash transactions, and online chat to allow individuals greater control over the exposure of their personal information.
- Well-designed software environments for writing Web pages have provided a low entry barrier to online publishing.

As a group, these exemplar projects have yielded new instruments (e.g., the Social Impact Statement), new process practices (e.g., documentation of processes as a part of

ISO 9001 registration), new design methods (e.g., for applying value analyses to the design of new technology), and awareness of economic considerations for Value-Sensitive Design (e.g., the economic ramifications for ignoring value considerations in design).

Note that these exemplar projects are closely aligned with the technical, human-computer interaction, and socio-economic research communities.

Section II: Research Agenda and Priorities

Participants developed working recommendations for a research agenda and priorities on Value-Sensitive Design. The major areas include theoretical and conceptual foundations, "proof-of-concept" projects, Value-Sensitive Design in context, criteria and metrics, and cultivating a Value-Sensitive Design research community. Each area is discussed below. Though each area is discussed separately, in practice there will often be an interplay, with individual research projects incorporating more than one area.

Theoretical and Conceptual Foundations

Value-Sensitive Design rests on the premise that technology is not value-neutral. That is, whether or not designers explicitly address values in their design process, it is not possible for designers to rid themselves of their values and design without them. In turn, the technology in some ways reflects designers' values. But, then, what does it mean to say that technology embodies human values? Or embeds them? Or provides suitabilities that are likely to support certain values over others? Building on the work of the Science-Technology-Studies community, this is a question for theoretical development. Moreover, we must engage in careful study of how to determine the values that are embedded in technology. We should not, for example, presume that software designers and engineers can make these decisions in isolation.

In order to address complex social values in the design of information technology that avoids caricature or naïve manifestation of those values, conceptual analyses of particular values are needed. We must study – systematically and comprehensively -- not only particular values in the online context, but also complexities that arise when trade-offs among competing values are required in a design (e.g., autonomy vs. security; anonymity vs. trust). Explicit connections need to be articulated between the value analyses and the realization of those values in an interactive context. Explicit work needs to be undertaken to understand how existing theories of design and software engineering processes can help to bridge the conceptual analyses with the practice of design and implementation.

Specific research recommendations include:

- Analyze how various technical mechanisms support specific value considerations.
- Conduct empirical and conceptual analyses of individual values and their ethical implications in the interactive context.
- Examine trade-offs among competing values in the design, implementation, and use of information systems.

- Examine how to translate a well-analyzed value into a technical implementation that respects that value.
- Develop technical means to assure that systems realize specific values as intended.
- Examine how values may differ in nature and priority across cultures and contexts.
- Analyze the relevance of theories of design for Value-Sensitive Design.
- Examine how Value-Sensitive Design can borrow from existing successful design methods.
- Examine where Value-Sensitive Design needs to develop new methodologies.
- Review the literature across a range of technical areas to identify existing work that fits broadly within Value-Sensitive Design.

"Proof-of-concept" Projects

Participants recognized the need to make connections between the theoretical and conceptual value analyses of stakeholders and the design process. Early on, much of this work will be exploratory and take the form of "proof-of-concept" projects. Projects that proactively apply Value-Sensitive Design theory and methods to particular designs can, at times, be complemented by retrospective studies of deployed technologies and designs including those that succeed in supporting value considerations and those that fall short on value dimensions.

Specific research recommendations include:

- Conduct "proof-of-concept" projects in which teams of researchers and designers apply the principles, theories, and methods of Value-Sensitive Design to a particular technology, domain, or design problem.
- Analyze "reasonably" successful real world technical mechanisms, products, and standards to gain insight into when and how technical mechanisms effectively support value considerations.
- Conduct retrospective studies of deployed technologies and designs that fall short on value dimensions. The goal here is twofold: to examine (1) what can go wrong in the design process, and (2) how to recover with minimal damages when designs fall short.

Value-Sensitive Design in Context

Participants recognized the contextual challenges to Value-Sensitive Design. Different stakeholders who influence the design and/or use a technology may have different goals and priorities that, in turn, lead to different value trade-offs. These differences may exist within a single organization or among organizations. Moreover, as environments change as the result of the infusion of technology, the trade-offs of values within and among organizations may shift. Thus, to be effective in practice, the general principles and methods of Value-Sensitive Design will need to be able to account for such diversity. Understanding how to integrate the theoretical foundations of Value-Sensitive Design with the complexities of design practice will require empirical study (of organizations, designers, technologists, and end-users).

Specific research recommendations include:

- Study how "typical" designers and software engineers appropriate and use Value-Sensitive Design methodologies, including espoused practice as compared with actual practice.
- Study how organizations appropriate Value-Sensitive Design, including their motivations, methods of training and dissemination, reward structures, and economic incentives.
- Examine how Value-Sensitive Design can be integrated with and augment existing design and software engineering practice.
- Examine organizational and individual obstacles to integrating value considerations into the design and software engineering process.
- Examine economic incentives and barriers to Value-Sensitive Design.

Criteria and Metrics

The workshop participants recognized the key role of well-defined criteria and metrics in clarifying goals and assessing progress for Value-Sensitive Design. Value-Sensitive Design criteria will help to focus research efforts and provide direction for incremental improvement. Once criteria are in place, corresponding metrics – qualitative and quantitative -- can be developed.

As an analogy, consider the criterion of reliability as a model: The criterion of reliability is an explicit ideal held out for the development of technical systems. Although we cannot build completely reliable systems, the criterion directs computer scientists to (a) develop methods that help to increase reliability, (b) establish metrics that provide a means to assess the reliability of a particular technology and to compare performance among systems, (c) create a shared language to refine a consensus on long-term goals and next-step research priorities, and (d) establish a shared understanding of what constitutes reasonable professional practice with respect to reliability. Taken together, these efforts allow computer scientists to discern which designs and which methods are effective. Similarly, well-articulated criteria and metrics are needed to firmly establish Value-Sensitive Design.

Specific research recommendations include:

- Develop Value-Sensitive Design criteria and corresponding metrics – qualitative and quantitative -- that can be used by system designers and others to guide the design process and assess the success of particular designs.
- Evaluate on-going designs in terms of the new Value-Sensitive Design criteria and metrics.

Cultivating a Value-Sensitive Design Research Community

To build on the recent interest in and momentum for Value-Sensitive Design, participants spoke to the importance of cultivating a research community. Developing such a research community serves many purposes, including (1) creating a critical mass of researchers who would further develop and deepen Value-Sensitive Design, (2) bringing new ideas and perspectives to the emerging work in Value-Sensitive Design, and (3) helping to infuse the goals and methods of Value-Sensitive Design into the work

of a broad community of researchers concerned with the design and deployment of information technology.

Specific recommendations include:

- Identify researchers for whom a Value-Sensitive Design perspective is critical to their existing research goals and endeavors (see Appendix 1 for an initial list).
- Train a new generation of researchers in the theories, research methods, and practices of Value-Sensitive Design. Such training would necessarily entail technical, social science, and values expertise.
- Train a new generation of practitioners (both in software engineering and more interaction-oriented areas) in the theories and practices of Value-Sensitive Design. How to conduct such training may require research into effective Value-Sensitive Design education.
- Organize (a) a small number of highly visible stand-alone workshops that bring together interested researchers across the relevant knowledge communities, (b) workshops attached to the professional meetings of relevant knowledge communities, and (c) tutorial sessions at professional conferences and industry training sessions.

References

Agre, P. E., & Harbs, C. A. (1994). Social choice about privacy: Intelligent vehicle-highways systems in the United States. *Information Technology and People, 7(4),* 63-90.

Baecker, R. M. (Ed.) (1992). *Readings in groupware and computer-supported cooperative work : Assisting human-human collaboration.* San Francisco, CA: Academic Press/Morgan Kaufmann.

Blomberg, J., Suchman, L., & Trigg, R. H. (1996). Reflections on a work-oriented design project. *Human-Computer Interaction, 11(3)*, 237-265.

Bodker, S. (1991). *Through the interface: A human activity approach to user interface design*. Hillsdale, NJ: Lawrence Erlbaum.

Clement, A., & Van den Besselaar, P. (1993). A retrospective look at PD projects. *Communications of the ACM, 36(4)*, 29-37.

Ehn, P. (1988). *Work-oriented design of computer artifacts*. Stockholm: Arbetslivscentrum. (Distributed by Lawrence Erlbaum, Hillsdale, NJ).

Ermann, M. D., & Williams, M. B., & Gutierrez, C. (Eds.) (1990). *Computers, ethics, and society*. Oxford: Oxford University Press.

Forester, T. (Ed.) (1989). *Computers in the human context*. Cambridge, MA: The MIT Press.

Friedman, B. (1996a). Value-sensitive design. *interactions*, III(6), 17-23.

Friedman, B. (Ed.) (1997a). *Human values and the design of computer technology*. New York: Cambridge University Press and CSLI, Stanford University.

Friedman, B., & Nissenbaum, H. (1996b). Bias in computer systems. *ACM Transactions on Information Systems, 14(3),* 330-347.

Friedman, B., & Nissenbaum, H. (1997b, February). Software agents and user autonomy. *Proceedings of the first international conference on autonomous agents* (p. 466-469). New York: Association for Computing Machinery.

Gotterbarn, D., Miller, K., & Rogerson, S. (1997, July). Software engineering code of Ethics. *SIGCAS Newsletter*.

Greenbaum, J., & Kyng, M. (1991). *Design at work*. Hillsdale, NJ: Lawrence Erlbaum.

Grief, I. (Ed.) (1988). *Computer-supported cooperative work: A book of readings*. San Mateo, CA: Morgan Kaufmann.

Grudin, J. (1999). *Information, context and conventions in computer-mediated interaction*. Manuscript submitted for publication.

Johnson, D. G. (1994). *Computer ethics* (2nd edition). Englewood Cliffs, NJ: Prentice Hall.

Johnson, D. G., & Mulvey, J. M. (1993). *Computer Decisions: Ethical Issues of Responsibility and Bias* (SOR-93-11). Statistics and Operations Research Series, Dept. of Civil Engineering and Operations Research, Princeton University, Princeton, NJ.

Johnson, D. G., & Nissenbaum, H. (Eds.) (1995). *Computers, ethics and social values*. Englewood Cliffs, NJ: Prentice Hall.

Kling, R. (1980). Social analyses of computing: Theoretical perspectives in recent empirical research. *Computing Surveys, 12(1),* 24-34.

Kling, R. (Ed.) (1996). *Computerization and controversy*: *Value conflicts and social choices* (2nd edition). Boston: Academic Press.

Kuhn, S. (1996). Design for people at work. In T. Winograd (Ed.), *Bringing design to software* (pp. 273-289). Reading, MA: Addison-Wesley.

Ladd, J. (1989). Computers and moral responsibility: A framework for ethical analysis. In C. Gould (Ed.), *The information web: Ethical and social implications of computer networking* (pp. 207-227). Boulder, CO: Westview Press.

Moor, J. H. (1985). What is computer ethics? *Metaphilosophy, 16*, 266-275.

Mumford, L. (1934). *Technics and civilization*. New York: Harcourt Brace and World.

Nardi, B. A., & O'Day, V. L. (1999). *Information ecologies: Using technology with heart*. Cambridge, MA: The MIT Press.

Nissenbaum, H. (1996). Accountability in a computerized society. *Science and Engineering Ethics, 2*, 25-42.

Nissenbaum, H. (1998a). Values in computer system design: Bias and autonomy. In *Ethics and information technology* (Proceedings of the Conference on Computer Ethics, Linkoping Sweden. June, 1997). Delhi: New Academic Press.

Nissenbaum, H. (1998b). Values in the design of computer systems. *Computers in Society*.

President's Information Technology Advisory Committee (PITAC). (1999, February 24). *Information technology research: Investing in our future* (Advance Copy). National Coordination Office for Computing, Information and Communications.

Schuler, D., & Namioka, A. (Eds.) (1993). *Participatory design: Principles and practices*. Hillsdale, NJ: Lawrence Erlbaum.

Shneiderman, B. (1991). Human values and the future of technology: A declaration of responsibility. *ACM SIGCHI Bulletin, 23(1)*, 11-16.

Shneiderman, B., & Rose, A. (1997). Social impact statements: Engaging public participation in information technology design. In B. Friedman, Ed., *Human values and the design of computer technology* pp. 117-133. New York, NY: Cambridge University Press.

Sproull, L., & Kiesler, S. (1991). *Connections: New ways of working in the networked organization.* Cambridge, MA: The MIT Press.

Suchman, L. (1987). *Plans and situated actions*. Cambridge: Cambridge University Press.

Suchman, L. (1994). Do categories have politics? The language/action perspective reconsidered. *CSCW Journal, 2(3)*, 177-190.

Tang, J. C. (1997). Eliminating a hardware switch: Weighing economics and values in a design decision. In B. Friedman, Ed., *Human values and the design of computer technology* pp. 259-269. New York, NY: Cambridge University Press.

Thomas, J. C. (1997). Steps toward universal access within a communications company. In B. Friedman, Ed., *Human values and the design of computer technology* pp. 271-287. New York, NY: Cambridge University Press.

Wiener, N. (1954). *The human use of human beings: Cybernetics and society.* Garden City, NY: Doubleday and Company.

Winograd, T. (1994). Categories, disciplines, and social coordination. *CSCW Journal, 2(3)*, 191-197.

Zuboff, S. (1988). *In the age of the smart machine*. New York: Basic Books.