Notre Dame Journal of Law, Ethics & Public Policy

Volume 14
Issue 2 *Symposium on Privacy and the Law*

Article 10

February 2014

Exploring Identity and Identification in Cyberspace

Oscar H. Gandy Jr.

EXPLORING IDENTITY AND IDENTIFICATION IN CYBERSPACE

OSCAR H. GANDY, JR.*

INTRODUCTION

This article is concerned with the relationship between identity and identification in the realm of cyberspace transactions. Philip Agre suggests that the commercialization of the Internet has increased the level of public concern about the ways in which personally identifiable information can be captured and used.[1] The demands of the marketplace have shaped and reflected the ways in which the capture of this information is automated and normalized. The marketplace may also have helped shape the consequences that flow from this normalization. This possibility is worthy of our concern.

Central to the discussion of the capture and use of personal information is the critical distinction between identity and identification. This distinction is based upon an elaborated notion of agency, or individual autonomy. Identity is associated with individual agency because it is related to the ability of the individual to shape her identity beyond the gaze and influence of powerful others. Those who advocate the defense of privacy against the threats of technology and social practice do so, in part, because of the belief that the "right to be left alone" is fundamental to the development of the autonomous individual.[2] Identification on the other hand, is understood in the context of the exercise of power and authority. Identification may be appreciated as an element of constraint within the process of structuration that Anthony Giddens has helped us comprehend.[3]

This article will also explore many of the tensions that have emerged within the evolving discourse on technology. Conflict is

* Herbert I. Schiller Information and Society Professor, Annenberg School for Communication University of Pennsylvania.

1. *See* Philip Agre, *The Architecture of Identity: Embedding Privacy in Market Institutions*, 2 COMM. & SOC'Y 1, 25 (1999).

2. Samuel Warren & Louis Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193-220 (1890); *see also* Katrin Byford, *Privacy in Cyberspace: Constructing a Model of Privacy for the Electronic Communications Environment*, 24 RUTGERS COMPUTER & TECH. L.J. 1 (1998).

3. *See* ANTHONY GIDDENS, THE CONSTITUTION OF SOCIETY: OUTLINE OF A THEORY OF STRUCTURATION (1984).

especially sharp with regard to the extent to which feasibility determines necessity. Jacques Ellul's engagement with "la technique" suggests that the availability of a technique not only invites, but demands its use.[4] Technological systems become defined as essential long before the claims made on their behalf have been weighed in the balance of private and social benefits and costs.[5] Widespread use contributes to the expectation of use.[6] Media coverage facilitates the normalization of a rationale of business necessity as a justification for expanded use. Business necessity is a primary justification for the use of the technologies of identification as adjuncts to traditional methods of drawing invidious distinctions among members of disfavored groups.[7]

This article is critical of business practices, as well as the rationales presented in their defense. It is especially critical of the justifications that are presented in support of the use of transaction-generated information in the development of predictive models that guide interactions with individuals in their roles as employees, citizens and consumers. At the core of this critique is a challenge to the use of statistical information as an aid to decision-making. The widespread view that statistical discrimination is both rational and efficient will be challenged on both technical and ethical grounds.[8] I will argue that the use of group averages as a basis for discrimination is always problematic from the perspective of those on the long side of any contested exchange.[9]

Although there is an extensive literature that has followed efforts to limit the impact of discrimination in the areas of employment,[10] education,[11] insurance,[12] housing,[13] law enforce-

4. *See* JACQUES ELLUL, THE TECHNOLOGICAL SOCIETY (John Wilkenson trans., 1964).

5. *See* GENE ROCHLIN, TRAPPED IN THE NET: THE UNANTICIPATED CONSEQUENCES OF COMPUTERIZATION (1997).

6. *See* DAN SCHILLER, DIGITAL CAPITALISM: NETWORKING THE GLOBAL MARKETING SYSTEM (1999).

7. *See* Oscar H. Gandy, Jr., *Legitimate Business Interest: No End in Sight? An Inquiry into the Status of Privacy in Cyberspace*, 1996 U. CHI. L. FORUM 77, 137.

8. *See* CASS R. SUNSTEIN, FREE MARKETS AND SOCIAL JUSTICE 151-66 (1997).

9. *See* Samuel Bowles & Herbert Gintis, *The Political Economy of Contested Exchange, in* RETHINKING POWER 196-224 (Thomas E. Wartenberg ed., 1992) (discussing a variety of circumstances in which there is a marked power imbalance between two actors. In the case of employment relations, the employer, who can select from a great many potential employees, is said to be on the "short side" of the market. The employee, who is constrained by the relatively small number of equivalent opportunities for employment, is said to be on the "long side," and thereby at a distinct disadvantage in negotiating contracts).

10. *See* Steven L. Willborn, *The Disparate Impact Model of Discrimination: Theory and Limits*, 34 AM. U. L. REV. 799, 836-37 (1985).

ment,[14] and even in the area of environmental justice,[15] the forms of discrimination especially likely to be encountered in the Internet environment are just beginning to attract scholarly attention. This article will explore the ways in which marketing discrimination[16] promises to introduce distortions into the sphere of e-commerce. Special attention will be paid to the ways in which these distortions will transform the market for information. The central role which access to information plays in the democratic process is the basis for special concern about this threat.[17]

The nature of discrimination against groups defined by race is familiar terrain. Yet, race is just one of the markers that are used to identify groups whose mistreatment demands our attention. We have seen the list of suspect categories expanded to include gender, age, and sexual orientation when they have been used as the basis for discrimination. This expansion has been based in part on the recognition that "differences from the community norm" can be the basis of invidious distinctions.[18] However, an expanding pattern of discrimination against non-racial groups[19] presents us with a difficulty in deciding which group is

11. *See* Vincent J. Roscigno, *Race and the Reproduction of Educational Disadvantage*, 76 SOC. FORCES 1033 (1998).

12. *See* Regina Austin, *The Insurance Classification Controversy*, 131 U. PA. L. REV. 517 (1983).

13. *See* Peter P. Swire, *The Persistent Problem of Lending Discrimination: A Law and Economics Analysis*, 73 TEX. L. REV. 787 (1995); *see also* Michael F. Ferguson & Stephen R. Peters, *What Constitutes Evidence of Discrimination in Lending?*, 50 J. FIN. 739, 748 (1995).

14. *See* SAMUEL WALKER ET AL., THE COLOR OF JUSTICE: RACE, ETHNICITY, AND CRIME IN AMERICA (1996).

15. *See* Lynn E. Blais, *Environmental Racism Reconsidered*, 75 N.C. L. REV. 75, 151 (1996).

16. *See* Timothy C. Lambert, Note, *Fair Marketing, Challenging Pre-Application Lending Practices*, 87 GEO. L.J. 2181 (1999).

17. *See* ROBERT W. MCCHESNEY, RICH MEDIA, POOR DEMOCRACY: COMMUNICATION POLITICS IN DUBIOUS TIMES (1999); *see also* Douglas Kellner, *New Technologies, the Welfare State, and the Prospects for Democratization*, *in* COMMUNICATION, CITIZENSHIP, AND SOCIAL POLICY: RETHINKING THE LIMITS OF THE WELFARE STATE 239-56 (Andrew Calabrese & Jean-Claude Burgelman eds., 1999).

18. Ian F. Haney-Lopez, *Race, Ethnicity, Erasure: The Salience of Race to LatCrit Theory*, 85 CAL. L. REV. 1143, 1203-05 (1997) (citing Hernandez v. Texas, 251 S.W.2d 531 (Tex. Crim. App. 1952), *rev'd*, 347 U.S. 475 (1954), where a Mexican American was denied the right to appeal discrimination under the Fourteenth Amendment because he was defined as White, and therefore not subject to racial discrimination).

19. *See* Oscar H. Gandy, Jr., *It's Discrimination Stupid!*, *in* RESISTING THE VIRTUAL LIFE: THE CULTURE AND POLITICS OF INFORMATION 35 (James Brooks & Ian Boal eds., 1995).

most deserving of protection.[20] The identification of these groups is made difficult by the need to consider the markers of group boundaries that are both visible and immutable, and those that are made visible only by means of statistical analysis.

I. IDENTITY

Identity formation is readily understood as a part of the process of individual development. The development of an identity distinct from that of one's parents is part of a tortured passage through adolescence.[21] Once formed, the identity of an individual is still complex and multidimensional. An individual's identity can be represented structurally as a multi-layered array. Near the base, we would expect to find the more stable foundations, perhaps reflective of racial, ethnic, gender, and perhaps political, national, or ideological identities. Higher, and moving out from the central core, we would find variations in the salience of particular aspects of identity reflecting the influence of situational cues.[22] The more central aspects of a person's identity function as resources and constraints that organize critical aspects of their daily lives.

It is important to note that individual identities are formed in interaction with others.[23] The characteristics of those interactions help to determine the salience, as well as the level of comfort with which different aspects of one's identity co-exist.[24] Self-esteem, or how an individual feels about herself is determined, in part, by the ways in which her relevant reference groups are evaluated by others.[25] The salience of racial and ethnic aspects of one's identity is determined by the interaction of a number of contextual factors. They include the distribution of political power and social status, the extent of residential segregation and occupational concentration, and the clarity with which racial or

20. *See* Charles Raab & Colin Bennett, *The Distribution of Privacy Risks: Who Needs Protection?*, 14 INFO. SOC'Y 263, 274 (1998).

21. *See* ERIK. H. ERICKSON, IDENTITY: YOUTH AND CRISIS (1968).

22. *See* HARRY H.L. KITANO, RACE RELATIONS 84-96 (1991).

23. *See* William E. Cross, Jr., *Oppositional Identity and African American Youth: Issues and Prospects, in* TOWARD A COMMON DESTINY: IMPROVING RACE AND ETHNIC RELATIONS IN AMERICA 185, 203 (Willis D. Hawley & Anthony W. Jackson eds., 1995).

24. *See* STEPHEN CORNELL & DOUGLASS HARTMANN, ETHNICITY AND RACE: MAKING IDENTITIES IN A CHANGING WORLD (1998).

25. Richard H. McAdams argues that racial discrimination and other hostile acts may be understood as attempts to build self-esteem by lowering the status of competing groups. *See* Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 HARV. L. REV. 1003, 1084 (1995).

ethnic distinctions are emphasized in routine interactions.[26] The mass media have also been identified as playing a critical role in the development of personal identity because of the ways in which media provide individuals with indirect access to the experiences of others.[27] A concern with the media's representation of racial and ethnic groups is based in part upon a belief that popular media provide representations that reinforce existing stereotypes. Those stereotypes help shape the ways in which people relate to each other.[28]

Because identity is formed through direct and mediated interaction with others, individuals are never free to develop precisely as they would wish. Parents, friends, and a host of authorities and interested parties actively seek to influence their evolution. On the basis of claims regarding the superiority of group interests over individual sensibilities, a host of public, and not so public figures have been called upon to expose or highlight one aspect of their identity more prominently than they might have wished.[29] Identification as a member of a group recognized in law may bring both benefits and harms.[30] Identifica-

---

26. *See id.* at 1090.

27. *See* DAVID CROTEAU & WILLIAM HOYNES, MEDIA/SOCIETY: INDUSTRIES, IMAGES, AND AUDIENCES 261-94 (2d ed. 2000) (discussing the concept of an "active audience" that interprets media representations. This is a view that rejects the more deterministic models of causality that emphasize the cultivation of social perceptions); *see also* Sandra J. Ball-Rokeach, *A Theory of Media Power and a Theory of Media Use: Different Stories, Questions, and Ways of Thinking*, 1 MASS COMM. & SOC'Y 5, 40 (1998).

28. *See* OSCAR H. GANDY, JR., COMMUNICATION AND RACE: A STRUCTURAL PERSPECTIVE (1998); *see also* JODY ARMOUR, NEGROPHOBIA AND REASONABLE RACISM: THE HIDDEN COST OF BEING BLACK IN AMERICA (1997) (discussing the influence of racial stereotypes on social interaction); C. Edwin Baker, *Giving the Audience What it Wants*, 58 OHIO ST. L.J. 311, 351, 366 (1997) (commenting on media externalities).

29. *See* PATRICIA BOLING, PRIVACY AND THE POLITICS OF INTIMATE LIFE 132 (1996) (exploring the strategic "outing" of prominent homosexuals. Although Boling suggests that the long term impact of forcing individuals "out of the closet" may be less than desirable, proponents argue that it is essential for the movement to be able to eliminate any connection between shame and homosexual identity).

30. Racial identification is a complex and uncertain process, due in no small part to the absence of any reliable authority. Its importance is increasingly linked to claims regarding rights and opportunities that have the character of licenses. *See, e.g.*, Michael Omi, *Racial Identification and the State: The Dilemma of Classification*, 15 LAW & INEQ. J. 7 (1997) (discussing racial and ethnic classification). With regard to racial identification as a form of property right, see Jim Chen, *Affirmative Action: Diversity of Opinions: Embryonic Thoughts on Racial Identity as New Property*, 68 U. COLO. L. REV. 1123 (1997); Cheryl I. Harris, *Whiteness as Property*, 106 HARV. L. REV. 1707 (1993) (discussing racial privilege as a property interest defended as a right); Alex M. Johnson, Jr., *Destabilizing*

tion *with* others may be the source of great pleasure and a sense of fulfillment. Identification *by* others is often experienced as an ungranted exercise of power.[31]

II. IDENTIFICATION

We understand identification as a process that varies in precision, and the demand for precision varies with the circumstance, or the nature of the interaction in which identification is an issue. We may think of identification as a continuum ranging from complete certainty at one end, to invisibility at the other.[32] Both certainty and invisibility are idealized and unattainable. At best, certainty might mean that the person identified through one procedure is the same person who would be identified by a large number of other independent means. This is the level of identification that is implied by the mathematical sign of equality. Practically, this means that the Davy Jones who stands before the clerk in the liquor store on New Years Eve, is the same Davy Jones who was born to Ralph and Elizabeth Jones in Bayonne, New Jersey at 4:27 in the morning on December 28, 1979. What matters in this particular circumstance, of course, is that Davy has provided some token of identification that indicates that he is at least twenty-one years of age, and therefore has the right to purchase a six-pack of alcoholic beverages. Neither his name, nor any other aspects of his identity are ordinarily considered to be relevant to this particular transaction.[33]

Not all interactions or transactions require the same level of identification. Indeed, most genuinely require no identification at all. Transactions that involve credit, or some other contractual assumptions about payment, or satisfactory performance, may require some form of reliable identification. However, the ease with which tokens of identity may be forged, or stolen, has raised

---

Racial Classifications Based on Insights Gleaned from Trademark Law, 84 CAL. L. REV. 887 (1996).

31. *See* RANDALL BARTLETT, ECONOMICS AND POWER: AN INQUIRY INTO HUMAN RELATIONS AND MARKETS (1989); Linz Audain, *Critical Cultural Law and Economics, the Culture of Deindividualization, the Paradox of Blackness*, 70 IND. L.J. 709 (1995).

32. Invisibility is used as an indication of the absence of identification because under that condition, no information about an individual is available as a basis for characterization.

33. *See* Roger Clarke, *Anonymous, Pseudonymous and Identified Transactions: The Spectrum of Choice* (extended abstract, visited Feb. 17, 2000) <http://www.anu.edu.au/people/Roger.Clarke/DV/UIPP99EA.html>; Roger Clarke, *When Do They Need to Know "Whodunnit?" The Justification for Transaction Identification; The Scope for Transaction Anonymity and Pseudonymity* (visited Mar. 20, 2000) <http://www.anu.edu.au/people/Roger.Clarke/DV/PaperCFP95.html>.

the general level of concern about authentication or verification of identity.

Some tokens require the actor or agent who accepts the token to participate in the process of authentication. Routines of verification may involve asking questions that the person presenting the token would be expected to know. Sophisticated tokens, such as those with embedded chips may reveal the authentication code only to the agent.[34] Transactions in cyberspace generally do not involve a human agent, and increasingly may be completed entirely by means of autonomous software agents.[35] Ordinarily, authentication will depend upon a variety of routines including passwords, and increasingly, biometric identifiers.[36] In the foreseeable future we can expect to see more personal computers delivered with cameras as well as microphones installed as a matter of course. The routine use of biometric techniques is likely to develop as a way of ensuring that the user has been reliably identified.[37]

Authentication in support of cyberspace transactions may require the participation of a trusted third party at some stage in the relationship between consumer and provider. The development of the technology, the routines, and the specification of the roles and responsibilities involved in the implementation of digital signatures represent an area of intense interest and uncertainty.[38]

---

34. *See* Roger Clarke, *Chip-Based ID: Promise and Peril* (visited Apr. 1, 2000) <http://www.anu.edu.au/people/Roger.Clarke/DV/IDCards97.html>.

35. *See* J.J. BORKING ET AL., INTELLIGENT SOFTWARE AGENTS AND PRIVACY: TURNING A PRIVACY THREAT INTO A PRIVACY PROTECTOR (1999).

36. *See* John D. Woodard, *Biometric Scanning, Law & Policy: Identifying the Concerns—Drafting the Biometric Blueprint*, 59 U. PITT. L. REV. 97 (1997).

37. IriScan, a corporation that specializes in the development and marketing of iris recognition technology has introduced a product for use with the personal computer. Its PC Iris product uses a hand-held personal iris imager that functions as a computer peripheral. According to its promotional material:

> After the PC Iris system is installed, the user simply holds the imager in his hand, looks into the camera lens from a distance of about 3"-4", and presses the start switch to initiate the identification process. If the presented iris matches a valid IrisCode for that system, privileges are granted within seconds.

IriScan Press Kit (visited Dec. 1998) <www.iriscan.com>.

38. *See* Adam White Scoville, *Clear Signatures, Obscure Signs*, 17 CARDOZO ARTS & ENT. L.J. 345 (1999); *see also* John C. Anderson & Michael L. Closen, *Document Authentication in Electronic Commerce: The Misleading Notary Public Analog for the Digital Signature Certification Authority*, 17 J. MARSHALL J. COMPUTER & INFO. L. 833 (1999); Kalama M. Lui-Kwan, *Electronic Commerce: Digital Signatures: Recent Developments in Digital Signature Legislation and Electronic Commerce*, 14 BERKELEY TECH. L.J. 463 (1999).

Reliable authentication helps to establish the requisite level of confidence that both parties to a transaction are *who* they say they are. While there is well-placed concern about fraudulent representation on the part of commercial vendors, or others seeking to take advantage of consumers,[39] our attention will be focused on the use of authentication, and identification in the context of what we would otherwise identify as legitimate transactions. Our concerns about invidious distinctions and patterns of discrimination are not activated at this level of identification. Those concerns begin to take shape when transaction-processing systems are charged with responsibility for knowing whether a given person is, or is not *what* she says she is. It is at this point that the assumptions of Peter Steiner's smug little pooch come under challenge.[40]

The use of tokens of identification and authentication routines as a means of determining whether the user is authorized to use the computer, or to access particular files may be quite reasonable.[41] Indeed, systems operators, like the vendors of alcoholic beverages may be required under the law to determine whether an individual is beyond some specified age, or does not reside in a jurisdiction that finds graphic sexual material beyond any redeemable social value.[42]

Providers of information services may reasonably seek to determine whether the visitor to the site is a registered subscriber. In similar fashion, systems operators may reasonably seek to determine whether visitors to interactive Web sites, such as chat rooms, newsgroups, forums, and Virtual Reality (VR) environments are qualified.[43] Membership in particular populations or social categories may be established as a qualification for entry.[44]

---

39. *See* John Rothchild, *Protecting the Digital Consumer: The Limits of Cyberspace Utopianism*, 74 IND. L.J. 893 (1999).

40. Peter Steiner's famous dog (cartoon) exclaims "On the Internet, nobody knows you're a dog." NEW YORKER, July 5, 1993, at 61.

41. Jerry Kang has proposed an operational standard of "functionally necessary use" that limits the use of personal data to that which is needed to complete a particular transaction. Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1249, 1293-94 (1998).

42. *See* Noah Levine, Note, *Establishing Legal Accountability for Anonymous Communication in Cyberspace*, 96 COLUM. L. REV. 1526 (1996).

43. *See* Chris Kramarae, *A Backstage Critique of Virtual Reality*, *in* CYBERSOCIETY 36, 56 (Steven G. Jones ed., 1995).

44. Membership agreements may require considerable detail about individuals. The criteria established for membership may exclude individuals on the basis of categorical distinctions, including those generally treated as suspect because they ask about age, gender, as well as race and ethnicity. *See* Philip

It has been argued, however, that the determination of the status of an individual as authorized to access information or to participate in a variety of interactions within electronic environments need not require any further identification of the individual. That is, once qualified, further transactions or interaction may be entirely anonymous, or pseudonymous.[45]

III. PROFILES

As Agre reminds us, there is no assumption that our desktop computer either knows, or cares who is using it.[46] However, it is in the nature of the Internet Protocol (IP) that personally identifiable information is made available for capture in every interaction between computers. The amount of personal information that is made available without the knowledge or consent varies as a function of the ways in which consumers are connected to the Internet. The Internet Service Provider (ISP) will require some form of identification, at the very least a user identification name, and a password. The user's IP address is personally identifiable information that is automatically exchanged in establishing the communications link. This address is generally stored by Web servers, and is frequently passed on to third parties for purposes of justifying claims for compensation by advertisers. The IP address contains information about the domain, the computer, the browser software, and a user account name that is likely to bear features in common with the user's own name.[47]

Of particular importance for the development of user profiles is the fact that each visit by the consumer to a Web site is recorded electronically. This information not only includes the date and time of each visit to specific resources on the server, but the Uniform Resource Locator (URL) of the referring server. Kang notes that when the reference is from a search engine, the keyword or search term that was used is also delivered.[48] A further complication is involved in the increasing use of tracking systems that deliver this user information to other servers without the user's knowledge.[49]

Giordano, *Invoking Law as a Basis for Identity in Cyberspace*, 1998 STAN. TECH. L. REV. 1, 79.

45. It is important, however, to distinguish between those types of interaction where the participants may remain anonymous to other participants, but may be uniquely, and reliably identified to the system operator or host.

46. *See generally* Agre, *supra* note 1.

47. *See* Kang, *supra* note 41, at 1224-30.

48. *See id.*

49. The ability of Web providers to insert hidden links into downloaded pages allows these links to direct users' browsers to sites that they may not have

Great concern has been expressed regarding the common practice of Web servers to identify individual users by storing an identifying code, a "cookie," that is stored on the consumer's hard drive until it is erased. Cookies facilitate the compilation of detailed records of a user's visits over an extended period of time. Additional techniques for delivering consumer's IP address to third party servers continue to be developed.[50]

Data mining is the term used to describe the technological approach to deriving strategic information from transaction-generated data. Whenever an individual engages in a transaction that utilizes a networked computer, a record of that transaction is generated, stored, and at some point, may be integrated into a dataset that can be mined for insights about underlying patterns of relationships. While transactions involving purchase and rental agreements are the primary sources of this increasingly valuable transaction data, even the search for information provides useful additions to this growing resource pool.[51] Because of improvements in the transmission of data, and dramatic reductions in the cost of storage and processing, the development of data warehouses invites the collection and integration of previously unrelated bits of information about consumers' transactions.

Data mining has a particularly troublesome aspect, as it relates to informed consent. Consumers can not provide truly informed consent regarding the use of transaction-generated information. Consumers can not possibly know how the information might ultimately be combined in order to produce some of those "valuable nuggets of information."[52] Those engaged in data mining have no predetermined hypotheses regarding the relationships they might uncover. "The data miner does not know, cannot know at the outset, what personal data will be of value or what relationships will emerge. Therefore, identifying a primary purpose at the beginning of the process, and then restricting one's use of the data to that purpose are the antithesis of a data mining exercise."[53] A variety of analytical techniques,

---

visited on their own. This feature is used to elevate the reported statistics of commercial sites, especially those that are involved in the provisions of sexual content. However, such unauthorized forwarding of client requests is also routinely used in the measurement of mainstream Web sites. *See* Kang, *supra* note 41, at 1228-30 (discussing these techniques).

50. *See generally* BORKING ET AL., *supra* note 35.

51. *See* Ann Cavoukian, *Data Mining: Staking a Claim on Your Privacy, Information and Privacy Commissioner/Ontario* (visited Feb. 26, 1999) <http://www.ipc.on.ca/web_site.eng/MATTERS/SUM_PAP/PAPERS/datamine.htm>.

52. *Id.* at 5.

53. *Id.* at 11.

including those that make use of neural networks and other techniques are relied upon to identify and define "associations, sequences, classifications, clusters" and forecasts.[54]

The most important products of data mining are the unique classifications of "types" of consumers, or information seekers, that can be characterized further by the addition of demographic information.[55] Members of the "Individual Reference Industry" claim that they limit their efforts to the most socially responsible activities. These activities include the prevention of fraud, consumer protection, locating organ donors, and providing assistance in child support cases.[56] Yet there are numerous other information vendors who specialize in providing individually identifiable information specifically for the purpose of enabling marketing discrimination.

We can understand the great variety in the sorts of consumer/citizen/person types that can be generated through data mining by considering the ways in which simple queries of geodemographic databases might be used.[57] A query of a geodemographic database would permit the identification of all the people within a given city, or county, or neighborhood whose income is likely to exceed some figure. They might also be identified on the basis of their ownership of particular automobiles, their purchase of gourmet foods, their membership in fraternal organizations, or the extent to which their preferences for action-adventure exceeds some criterion level.[58]

---

54. *Id.* at 6.

55. It is important to emphasize the relative ease with which individually identifiable information about a broad range of interests and activities can be acquired legally within the information marketplace. *See, e.g.*, Andrew L. Shapiro, *Privacy For Sale: Peddling Data on the Internet*, THE NATION, June 23, 1997, at 16; Rajiv Chandrasekaran, *Doors Fling Open to Public Records* (visited Apr. 12, 2000) <http://www.washingtonpost.com/wp-srv/frompost/march98/privacy9.htm>.

56. *See generally* Piper & Marbury L.L.P., *White Paper: Individual Reference Services* (visited Apr. 1, 2000) <http://www.search3.knowx.com/home/home.exe?form=white_paper1.htm>.

57. *See* John Goss, *Marketing the New Marketing: The Strategic Discourse of Geodemographic Information Systems, in* GROUND TRUTH: THE SOCIAL IMPLICATIONS OF GEOGRAPHIC INFORMATION SYSTEMS 130, 170 (John Pickles ed., 1995) [hereinafter GROUND TRUTH]; *see also* Michael Curry, *Geographic Information Systems and the Inevitability of Ethical Inconsistency, in* GROUND TRUTH, *supra*, at 68, 129; John Goss, *We Know Who You Are and We Know Where You Live: The Instrumental Rationality of Geodemographic Systems*, 71 ECON. GEO. 171 (1995).

58. *See, e.g.*, ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE, GETTING TO KNOW DESKTOP GIS (1995) (providing examples of the kinds of queries described in an introduction to a popular desktop Geographic Information System (GIS), ArcView). This introduction discusses the great variety of data that can be purchased in the market (and acquired online) that includes informa-

Providers of Web-based information services have increasingly turned towards advertising as a means of survival within an environment in which the dominant culture is hostile toward the idea of paying for information.[59] Advertising in the context of the World Wide Web is valuable to the extent that it can be targeted to categories of individuals who have by their recent behavior, "self-identified" as a member of a target group.[60]

Conflicts emerge most critically in the context of Internet service providers, such as American Online, that have a special relationship with their member/subscribers. Because these portals provide a broad range of interactive services that vary substantially in terms of the economic value and the sensitivity of the information that can be captured and stored by the provider during the ordinary course of business. These providers will almost certainly make use of transaction-generated information to improve their service to their members,[61] and we would assume

---

tion about the age, sex, race, income, as well their demand for a variety of health care services. ESRI demonstrates the usefulness of ArcView for the targeting of particular consumers through a fictional example. "Michael," seeking to determine the best location for his "Wild Outdoors" store, makes use of ArcView's "Query Builder" to identify those census tracts where at least 10 percent of the households would be members of four segments that he believes to be his most likely customers. The most likely and valuable customers are those in the segment identified as "Movers and Shakers," although Michael is willing to risk investing in the segment identified as "Successful Singles." *Id.* at 103-20. The ArcView software, in combination with the consumer data available from an army of vendors, makes it relatively easy for individuals to engage in geodemograhic segmentation and targeting. The specialized techniques developed by Jonathan Robbin, and marketed widely through his Claritas Corporation, have become an off-the-shelf aid to marketplace discrimination. For the history of Geodemographic clustering, see MICHAEL J. WEISS, THE CLUSTERING OF AMERICA (1988).

59. *See* HOWARD RHEINGOLD, THE VIRTUAL COMMUNITY 65, 109 (1993). This is a concern that emerges repeatedly in the context of discussions about copyright management technology. *See* Julie E. Cohen, *A Right to Read Anonymously: A Closer Look at Copyright Management in Cyberspace*, 28 CONN. L. REV. 981, 1039 (1996); Malla Pollack, *The Right to Know?: Delimiting Database Protection at the Juncture of the Commerce Clause, the Intellectual Property Clause and the First Amendment*, 17 CARDOZO ARTS & ENT. L.J. 47 (1999).

60. *See generally* Sharon Machlis, *Net Bids for Ads Tests Privacy: Deja News Develops Plans to Better Target Advertising on its Web Site*, COMPUTERWORLD, June 29, 1998, at 3.

61. *See generally* Ellen M. Kirsh et al., *Recommendations for the Evolution* of Cyberlaw, 2 J. COMPUTER-MEDIATED COMM. (visited June 25, 1999) <http://www.ascusc.org/jcmc/vol2/issue2/kirsh.html>. Kirsh and her colleagues argue that interactive service providers "need" to provide their independent content providers, their advertisers, merchants and other partners with "aggregate information" that can be used for editorial, programming, and demographic analysis. They suggest that the absence of reliable audience data keeps elec-

that they would avoid those uses that might generate substantial doubt or mistrust regarding the use of personal information.[62] Yet, the tempting appeals from advertisers to supply more and more detailed information about member/subscribers may be irresistible.

Robin Mansell and her colleagues at the University of Sussex have explored the problems that commercial developers of "virtual communities" are facing as the potential for secondary exploitation of the information generated within these communities increases.[63] Among the issues that have emerged are many which have to do with the nature of identity and identification. Within the "gift economies" that are common to virtual communities, there is genuine value in the identities that have been formed through interaction within the community over time.[64] Their identity within the group bears no necessary relationship to any identity, or systems of identification that operate outside the virtual community. Yet, for a number of legal and financial reasons, the provider, or entities in partnership with the Internet Service Provider (ISP), may wish to link virtual identities with identities in the "real world."

Virtual communities and groups may establish their own rules about the nature and extent of identification within a particular locale that may differ from the more general rules established by the service provider. The rules may vary as a function of the sorts of transactions or interactions most likely to take

---

tronic publishers from garnering revenue from advertisers. Indeed they suggest that armed with transaction-generated data, providers will be able to "present the user with information she might want or need before she even realizes it exists." *Id.* Denying providers access to the data to meet these legitimate needs would keep them from serving their customers. Kirsh and her colleagues were attorneys employed by American Online when this article was published.

62. Of course, service providers are repeatedly surprised by the hostile reaction of consumers to what they saw as innovative, harmless uses of transaction-generated information. The response of IBM's chairman, Louis Gerstner, to the release of information about what books IBM employees were buying from Amazon.com was anything but pleasant, and was accompanied by the threat of a corporate boycott. *See* Ira Sager, *How Amazon Got on IBM's Bad Side*, BUS. WK., Sept. 27, 1999, at 6; *see also* Jeri Clausing, *Intel Alters Plan Said to Undermine PC Users' Privacy* (visited Apr. 3, 2000) <http://www.nytimes.com/library/tech/99/01/cyber/articles/26internet.html>.

63. ROBIN MANSELL ET AL., SCIENCE AND TECHNOLOGY POLICY RESEARCH, UNIVERSITY OF SUSSEX, NET COMPATIBLE. VIRTUAL COMMUNITIES, INTELLIGENT AGENTS AND TRUST SERVICE PROVISION FOR ELECTRONIC COMMERCE (1998).

64. The value of one's reputation as a source of useful information, good humor, or common sense is an asset that is built up over time, and may be linked to a purely pseudonymous identity within the context of a virtual community such as a MUD, or MOO, or even a newsgroup.

place within those cyberspaces. Jerry Kang has identified three different sorts of spaces in which rules regarding racial identification might be established.[65] Those spaces identified most readily with commercial transactions would be characterized by the absence of racial identifiers. Those spaces in which non-economic social interaction is more likely, authenticated racial identification might be required or facilitated. A third variety of cyber-spaces would include those in which identity play is common and where racial identification may be presented, but not authenticated.

The problems involved in facilitating the transfer of identity between cyberspaces without violating the rules governing the nature of identification are substantial.[66] Yet, they pale in comparison with the temptation to acquire, combine, and interpret the information generated through interactions and transactions in these spheres. The fact that a large number of the most highly visible portal and virtual communities, and transaction service providers make use of intelligent agents and other techniques to capture transaction generated information is consistent with this view.[67]

Problems with maintaining a trusted relationship with users have led service providers to establish personalized or "customized" points of entry which gather personal information at the time of registration, and accumulate additional information through transactions that are enabled through those personalized pages. Although they have not been widely implemented, it

65. Jerry Kang, *Cyber-Race*, 113 HARV. L. REV. 1130 (2000).

66. Of course, there is also the continuing need for protecting individuals from those who would engage in deception, or hostile action requiring the use of force, or police powers in the real world. *See* Rothchild, *supra* note 39 (describing deceptive marketing). For a discussion of the problems of the development and enforcement of community standards within an electronic community, see Julian Dibbell, *Taboo, Consensus, and the Challenge of Democracy in an Electronic Forum*, *in* COMPUTERIZATION AND CONTROVERSY 552, 568 (Rob Kling ed., 2d ed. 1996).

67. Robin Mansell and her colleagues interviewed top tier portals like Yahoo, that began as search engines or directories, second tier sites that support virtual communities, including Tripod, as well as those like Auto-by-Tel that provide referral services. *See* MANSELL, *supra* note 63. *See also* Noah Robischon, *Browser Beware*, BRILL'S CONTENT MAG., July 1998, at 40, 44 (discussing the ways in which Yahoo, Excite, and other providers have moved from information providers to marketing resource). Robischon emphasizes the fact that searches that utilize the "channels" provided by Yahoo and Excite give preferential treatment to advertisers who are clients. *See id.*

seems likely that some form of individualization with regard to information collection and use will become standard.[68]

IV. DISCRIMINATION

The computer profile is a discriminatory technology. It is a resource used to differentiate between persons and groups. We are concerned about these distinctions because they nearly always involve evaluation.[69]

While the popular perception of a profile is linked to a dossier that contains exquisite detail about an individual, a cyberspace profile is something different. A profile is primarily a list of categories that have been determined to be relevant to some administrative decision that must be made by an organization with regard to an individual, a group, or a class. Individual categories or variables are the dimensions along which an entity may be evaluated. Subsets of categories may be combined into an index score. The fundamental purpose of a profile is the assignment of an individual into a class or category that represents a decision. This is a process of identification with a consequence.

There are at least three bases for concern regarding the use of transaction-generated information for the classification of individuals (profiling) for use in decisions that may determine economic opportunity. First, the data may be erroneous or incomplete. Second, the models used to interpret the relations

68. The Platform for Privacy Preferences (P3P) represents one technical response to the problem of negotiating agreements about the collection and use of transaction-generated information. The technology, still under development, would allow Web sites to identify their privacy practices, at the same time that users can express their own preferences regarding the collection of use of information generated by their use of the service. Optimally, the agent software would be able to negotiate agreements on the fly, perhaps requiring active consideration by the user only when standard options are not available. As currently being developed, agreements might involve a matrix of sensitivity defined by the intersection of data categories and data practices. The most sensitive might be demographic and preference data that could be disclosed to others for marketing purposes. *See* Lorrie Faith Cranor & Joseph Reagle, Jr., *Designing a Social Protocol: Lessons Learned from the Platform for Privacy Preferences Project, Apr. 1998 version* (visited Feb. 26, 2000) <http://www.w3.org/People/Reagle/papers/tprc97/tprc-f2m3.html>.

69. Evaluation seems to be a fundamental aspect of the meanings we assign to concepts and things. Repeated factor analysis of semantic differential scale items consistently produce three dimensions: evaluation, potency, and activity. Different concepts, such as father and mother can be located in multidimensional semantic space as a function of their differential weights on these dimensions. I interpret this widespread tendency as evidence that difference implies evaluation. *See* Donald K. Darnell, *Semantic Differentiation, in* METHODS OF RESEARCH IN COMMUNICATION 181, 196 (1998).

between variables may be flawed. Third, the consequences that flow from the use of the models may exacerbate existing disparities.[70]

Data may be in error for a variety of reasons. These include the existence of biases in the systems that capture, process, and report data as facts.[71] That models may be in error hardly needs statement. The fact that models are resistant to challenge, and are little threatened by disconfirming evidence,[72] is even more worrisome when we recognize that these models are incorporated into the routine assessments that are made by computers.[73] Another shortcoming that operates at the level of both theory and data is the assumption that the identity of an individual can be reduced, captured, or represented by measurable characteristics. The fact that the most easily obtainable measure is that which marks decisions about consumption suggests that the identities that can be constructed are unidimensional.[74] Goss charges

---

70. *See* Barbara D. Underwood, *Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgement*, 88 YALE L.J. 1408 (1979) (providing a comprehensive review and assessment of the arguments against the use of predictive models to guide decisions about individuals). Professor Underwood argues most convincingly in support of the demand that more than simple correlation be the basis for a model's use. A causal theory is a basic requirement, and for that theory to be useful, it must incorporate an appreciation of the options available to those about whom decisions will be made. *See id.* at 1448.

71. We understand information generally as reflecting a set processes that involve abstractions that introduce variation, if not error in the representation of facts in the real world. Robert M. Hayes describes this chain as a path that moves from facts, through their representation in data, through their processing into information, that might be communicated in ways that generate understanding, that would become integrated within a comprehensive system we would recognize as knowledge, and would be applied in the context of ethical decision making that we might recognize as having been wise. *See* Robert M. Hayes, *Measurement of Information Communication: A Set of Definitions, in* BETWEEN COMMUNICATION & INFORMATION 81, 103 (Jorge R. Schement & Brent D. Ruben eds., 1993).

72. Melvin W. Reder discusses the failure of contrary empirical evidence to weaken the commitment of economists to the "rational actor paradigm." MELVIN W. REDER, ECONOMICS: THE CULTURE OF A CONTROVERSIAL SCIENCE 160 (1999).

73. *See* Brian Cantwell Smith, *Limits of Correctness, in* COMPUTERS IN COMPUTERIZATION AND CONTROVERSY 810, 825 (Rob Kling ed., 1996).

74. It is a common, but nevertheless important criticism of predictive models. That which is easily measured is most likely to be included in the model. The circumstances that determine which features of a complex social environment will be subject to measurement are often ignored in the evaluation of a model's specification. *See* SUNSTEIN, *supra* note 8, at 128; *see also* WILLIAM H. DUTTON & KENNETH L. KRAEMER, MODELING AS NEGOTIATING: THE POLITICAL DYNAMICS OF COMPUTER MODELS IN THE POLICY PROCESS (1985).

geodemographic analyses with being guilty of the "fetishization of life-style" or presenting consumer life-styles "as if they were independent of socioeconomic relations."[75]

The use of predictive models based on historical data is inherently conservative. Their use tends to reproduce and reinforce assessments and decisions made in the past. Examples abound. Even though Cass Sunstein has suggested that statistical discrimination may be economically rational,[76] he suggests that its social impact may be socially destructive because of the effect that such decisions may have on the development of human capital. Those not yet in the market may observe the operation of the market, and observe the meager returns to investment in education and training. They may, as rational actors, decide that it makes little sense to invest in education, when other activities less beneficial to society as a whole may provide more immediate rewards. Of course, the failure to invest in human capital, and then to pursue disfavored alternatives for survival, only serve to reinforce stereotypic views. The vicious circle is drawn tighter still.[77]

V. SEGMENTATION

The development and use of consumer profiles facilitates the use of market segmentation as a basis for increasing sales and profits. Market segmentation is pursued along an almost unlimited number of dimensions.[78] The social consequences that flow from the use of market segmentation are seen as problematic by many contemporary observers.[79] Our concerns with regard to the social consequences of segmentation tend to be focused on segments based on demographic categories, including race and ethnicity.[80] Segmentation and targeting on the basis of revealed preferences has not been subject to the same level of critical review. Lambert has argued that market segmentation is a form a marketing discrimination. He suggests that "[t]he best terms

75. JOHN GOSS, WE KNOW WHO YOU ARE 50 (1995).

76. Not only does he suggest that categorical behavior informed by stereotypes may be economically rational as a basis for making employment decisions, he goes on to suggest that such decisions are "entirely legitimate in most settings." SUNSTEIN, *supra* note 8, at 151, 155.

77. *See id.* at 155, 162.

78. *See generally* JOSEPH TUROW, BREAKING UP AMERICA: ADVERTISERS AND THE NEW MEDIA WORLD (1997).

79. *See id. See also* Elihu Katz, *And Deliver Us from Segmentation*, 546 ANNALS, AAPSS 22, 33 (1996).

80. *See generally* OSCAR H. GANDY, JR., RACE, ETHNICITY AND THE SEGMENTATION OF MEDIA MARKETS IN MASS MEDIA AND SOCIETY (James Curran & Michael Gurevitch eds., 3d ed. Forthcoming 2000).

and conditions for credit cards, home equity loans, and other forms of credit may now elude minorities not because a loan officer turned them down, but because they never received an offer in the first place."[81] This process becomes especially troublesome in the context of an emerging tendency to send pre-approved applications to selected consumers. He suggests that unlawful discrimination might also be charged if the distribution of offers reflect the identification of a neighborhood as being ineligible on the basis of race. He suggests further that a successful claim may also be made on the basis of advertising that avoids minority media outlets.[82]

Many commentators emphasize the impact of segmentation on the operation of the public sphere.[83] C. Edwin Baker discusses the ways in which an advertiser-supported press may be involved in the development of an inauthentic or "corrupt" segmentation. Authentic segmentation would be related to "each group's discursive development within the lifeworld in response to each group's identification of its needs and values."[84] Corrupt segmentation would, conversely, reflect the influence of money, rather than group needs or values. Baker agrees that "[m]arket steering can equally corrupt segmentation" and that "this corrupt segmentation undermines both common discourse and self-governing group life."[85]

The social consequences that flow from consumer segmentation targeting are largely invisible to the participants in the contemporary debates about privacy and regulation of Internet commerce because the attention of both scholars and policy makers has been focused on the expressions of concern by individuals.[86]

VI. PRIVACY POLICY

A focus on the concern of individuals about their own vulnerability is understandable. Even when the right of privacy is

81. Timothy C. Lambert, *Fair Marketing: Challenging Pre-Application Lending Practices*, 87 GEO. L.J. 2182, 2183 (1999).

82. *See id.* at 2208.

83. *See, e.g.*, Oscar H. Gandy, Jr., Dividing Practices: Segmentation and Targeting in the Emerging Public Sphere (June 1998) (unpublished conference paper, on file with author).

84. C. Edwin Baker, *The Media That Citizens Need*, 147 U. PA. L. REV. 317, 376 (1998).

85. *Id.* at 377.

86. *See* MARY CULNAN & SANDRA MILBERG, CONSUMER PRIVACY IN INFORMATION PRIVACY: LOOKING FORWARD, LOOKING BACK (Mary Culnan et al. eds., forthcoming 2000)

examined in the context of group rights, that right is understood to have been derived from individual rights and from the benefits of association.[87]

One may argue that there is an intrinsic value in the group as a social collective. This value is to be seen as part of a continuum of existence that flows from personhood, through communality, to sociality. There may be a category of rights appropriate for each.[88] Garet challenges our assumptions about the basis for our appreciation of groups. He argues that "[w]hile some groups of importance in our lives are voluntary associations, others, of perhaps greater moral bearing, are neither voluntary nor associative."[89] Instead, Garet argues, the Supreme Court has assigned great value to the rights of groups that are "highly ascriptive; membership is imputed rather than chosen in any ordinary sense."[90] While I cannot yet claim that the sorts of groups that are created by means of their common victimization will eventually acquire a sense of solidarity and common purpose in opposition, their existence and their group status cannot be denied.

Privacy legislation has traditionally been concerned with individually identifiable information. Information about groups, or aggregates, is seen to be beyond the reach of privacy's scope. Indeed, it has been suggested that because aggregate information does not identify persons individually, there is no threat to privacy. Even the most expansive efforts to limit the commercial use of consumer information specify exceptions for aggregate information.[91] This does not mean that social groups, such as

87. *See* Edward J. Blaustein, *Group Privacy: The Right to Huddle*, 8 RUT.-CAM. L.J. 219, 222 (1977).

88. *See* Ronald R. Garet, *Communality and Existence: The Rights of Groups*, 56 S. CAL. L. REV. 1001, 1075 (1983). Garet links our interest in personhood, sociality, and communality with the ideals of the French Revolution in the call for "Liberté, Egalité, Fraternité." *Id.* at 1074.

89. *Id.* at 1043 (referring to the religious community in *Wisconsin v. Yoder*, 406 U.S. 205 (1972), and the tribe in *Santa Clara Pueblo v. Martinez*, 436 U.S. 49 (1978)).

90. *Id.* at 1046.

91. The Cable Communications Policy Act of 1984, 47 U.S.C. § 551 (a)2 (A) (1995) explicitly excludes "any record of aggregate data which does not identify particular persons." The Telecommunications Act of 1996, 47 U.S.C. § 222 (c)(3) (1998), indicates that a telecommunications carrier "may use, disclose, or permit access to aggregate customer information," and more pointedly, section (f)(2) defines aggregate customer information as "collective data that relates to a group or category of services or customers, from which individual customer identities and characteristics have been removed." 47 U.S.C. § 222 (f)(2) (1998).

women, have not been involved in presenting arguments about the ways in which privacy laws affect their interests.[92]

This does mean, however, that groups whose existence is entirely conceptual, in that they are unknown to each other, have little possibility of acting as a group. They are also unlikely to realize the benefits of taking legal action as an aggrieved class or otherwise shaping the privacy debate. Aggrieved individuals, believing that others share their experience, may take the initiative, and may even use the power of the Internet to identify others and begin the process of interest group formation. Enterprising lawyer advocates may also assume the responsibility for developing a plaintiff class and helping to shape its approach to litigation.[93] Whatever the process, the legitimate interest of these groups in being spared the costs of invidious discrimination deserves consideration.

A. *The Velvet Trap*

The dominant trend within the debates surrounding privacy and the emerging Internet is an embrace of the market as a solution to the problems of corporate abuse. Self-regulation is the form that a marketplace solution would ultimately take. It is argued that rational self-interest would govern choices made by firms seeking to succeed in e-commerce. Competition is supposed to ensure that firms cannot survive and prosper if they engage in activities that consumers find objectionable. Yet, we know that this idealized competitive market is rarely seen in the world of flesh and blood,[94] and is hard to imagine in cyberspace.

Jerry Kang examines the marketplace solution and suggests that the strongest argument against it is one that is based on a concern for the autonomous individual.[95] Part of this critique suggests that consent within this market is not voluntary, but coerced. Consent in this market is also uninformed because of the informational asymmetry that characterizes most market transactions. Not only are consumers unaware of the nature of

92. The claims by feminists that privacy rights have served to create and perpetuate women's unequal status certainly qualifies as an example. *See* Linda C. McClain, *Reconstructive Tasks for a Liberal Femnist Conception of Privacy*, 40 WM. & MARY L. REV. 759, 793 (1999). *See also* PATRICIA BOLING, PRIVACY AND THE POLITICS OF INTIMATE LIFE 132-56 (1996) (discussing issues related to gay and lesbian identity and the practice of outing).

93. *See generally* Ann Southworth, *Collective Representation for the Disadvantaged: Variations in Problems of Accountability*, 67 FORDHAM L. REV. 2449 (1999).

94. *See* ROBERT KUTTNER, EVERYTHING FOR SALE: THE VIRTUES AND LIMITS OF MARKETS (1999).

95. *See* Kang, *supra* note 41, at 1265-66.

the informational dimensions of their transactions, they are especially unaware of the consequences that flow from the use of their transaction-generated information in the development of profiles. Indeed, I would argue that they are largely unaware that they are even participating in what is being discussed as a "second exchange," where some "intangible benefits" like "higher quality service" and "personalized offers" are received in exchange for the right to use personal information.[96] As Culnan and Milberg note, because Web servers capture information about browsing, as well as information about actual transactions, this so-called "second exchange" often takes place before, and in many cases, without the primary exchange ever taking place.[97]

B. *First Principles*

Kang offers recommendations for the establishment of particular default rules that would govern the flow of personal information in the absence of explicit contracts governing transactions. He bases his recommendations on the concept of functional necessity. He determines that the "processing of personal information for any form of advertising—even when that advertising is done by the information collector—is not functionally necessary. Disclosing personal information to third parties to do the same would, *a fortiori*, not be functionally necessary."[98] We have little reason to expect, or even to hope that Kang's default rules might actually reach this fundamental limitation. Instead, we have to look for any signs that the policy system is even open to the possibility of establishing limits on the power of business to exploit identity and identification without regard to the consequences.

The Federal Trade Commission (FTC) clearly reflects the view that self-regulation is preferable to the more active regulatory stance favored by members of the European Community.[99] The Commission's recent assessment of industry practice has

96. *See* Mary J. Culnan & Sandra Milberg, *Consumer Privacy* (visited Apr. 14, 1999) <http://www.msb.edu/faculty/culnanm/research/conspriv.pdf>.

97. *Id.* at 14. Other observers note that many consumers make use of the Internet to gather information about products that they will purchase in neighborhood stores because of the benefits associated with being able to return and exchange goods.

98. Kang, *supra* note 41, at 1272.

99. *See* FEDERAL TRADE COMMISSION, SELF-REGULATION AND PRIVACY ONLINE: A REPORT TO CONGRESS (visited Sept. 1, 1999) <http://www.ftc.gov/os/1999/9907/privacy99.pdf> [hereinafter FTC]; *see also* Mark E. Budnitz, *Privacy Protection for Consumer Transactions in Electronic Commerce: Why Self-Regulation is Inadequate*, 49 S.C. L. REV. 847 (1998); Joel R. Reidenberg, *Restoring Americans' Privacy in Electronic Commerce*, 14 BERKELEY TECH. L.J. 771 (1999).

been focused on the extent to which the information practices of the most popular commercial Web sites implement each of "four substantive fair information practice principles" that have emerged as the de facto standard in the United States.

The four principles include the most fundamental—that consumers should be informed about information practices before the information is collected.[100] Once informed, a second core principle makes sense—that consumers be given the opportunity to choose options based upon their understanding of how those options might govern the collection and use of personal information.

The third principle would ensure that consumers have "reasonable" access to the information about them that has been collected, as well as an opportunity to "contest" the data's accuracy and completeness. Finally, consumers would be assured that their information would be secure from unauthorized use. The Commission concluded that while some progress had been made since their 1998 report, few of the busiest Web sites had implemented all four principles to the Commission's satisfaction.[101] Yet, on the basis of the evidence available to them, the Commission concluded that it would not be appropriate for the government to pursue a legislative or regulatory solution at this time.

Neither the FTC nor the Department of Commerce appear likely to include the "Use Limitation Principle" in any Codes of fair information practices that will be used in evaluating the Internet marketplace. The Use Limitation Principal, which is a close correlate of the principle that the use of information should be fully specified at the time of collection, suggests that personal data should not be disclosed, made available or otherwise used for purposes other than those specified initially.[102] The development of consumer profiles, especially those that are developed by means of data mining operations that combine information from a broad range of transactions, cannot be understood as being consistent with this principle. Data mining

100. What matters of course, is that consumers understand the ways in which transaction generated information is used in the creation of profiles that are then used in determining their access to opportunities, and advantages in future transactions, or in unexpected barriers to other important opportunities. The questions proposed for consideration during the workshop revealed a considerable awareness of the nature of profiling, but a lack of understanding of the relationship between organizations that produce profiles, and their clients.

101. *See* FTC, *supra* note 99, at 6.

102. *See OECD Privacy Guidelines, in* THE PRIVACY LAW SOURCEBOOK 181 (Marc Rotenberg ed., 1999).

is, by definition, a secondary use.[103] And under these international principles, "requires the explicit consent of the data subject or consumer."[104]

It is also abundantly clear that data mining is not, and is not very likely to ever be an open and transparent activity. Privacy Commissioner Ann Cavoukian suggests: "Even consumers with a heightened sense of privacy about the use and circulation of their personal information would have no idea that the information that they provided for the rental of a movie or a credit card transaction could be mined and a detailed profile of their preferences developed."[105]

It is also not clear whether the organizations making use of data mining technology would be prepared to reveal to consumers the nature of their profiles, as the FTC code suggests.[106] Nor is it clear that an individual consumer should be expected to provide an informed and compelling challenge to any aspect of their individual profiles. Literally thousands of transactions by one person are interpreted in the context of thousands more made by many thousands of unidentified others.[107] This complexity is simply not open to challenge, even if it were presented to consumers in some intelligible form.

The fact that the FTC, in cooperation with the Department of Commerce, announced a workshop focusing on online profiling and reported plans for an investigation of the use of electronic identifiers to track consumer behavior is encouraging. Yet, there are few signs that the Commission, or any other federal agency appears prepared to engage the hard questions about the ways in which transaction-generated information is used to identify the groups which will be the targets of market segmentation and other forms of discrimination.[108]

---

103. *See* Robert O. Harrow, Jr., *Are Data Firms Getting Too Personal?* (visited May 4, 1999) <http://www.washingtonpost.com/wp-srv/frompost/march98/privacy8.htm>. Harrow notes that the number of firms engaged in data mining numbers in the thousands, up from a few hundred in the recent past.

104. Cavoukian, *supra* note 51, at 10

105. *Id.* at 11.

106. *See* Harrow, *supra* note 103 (identifying Acxiom Corp. as an industry leader in data mining and supplier of consumer information to advertisers that argues that it would be impractical to provide individuals with access to their files).

107. *See* H. Jefferson Smity, *Do Data Warehouses Challenge Fair Play? Computers and Society* (visited Feb. 16, 1998) <http://www.ibm.com/OtherVoices/BeyondComputing/May2397185954.phtml>.

108. It will be essential for the government agencies responsible for the oversight of cyberspace transactions to understand the ways in which discrimination by profile actually works. While organizations that mail special offers, or

If we pursue the interest of those groups who are the victims of invidious distinction and marketing discrimination, we might ask which group should we protect first?

VII. In Defense of Invisible Groups

Charles Raab and Colin Bennett have asked about the distribution of privacy risks.[109] They explicitly consider the possibility that policy actors may need to identify "particularly vulnerable social groups" for the purpose of building coalitions and mobilizing advocates who support their specific interests, as well as those organizations concerned with civil liberties more generally. They note that policy discourse that is concerned with "fair information principles" is generally not concerned with achieving equality. They note as well, that in the United States, the general practice is one that leaves "it largely to individuals themselves to pursue complaints and to seek remedies."[110] The policy problem is not appreciably less troublesome in the European context. There, we observe a concern with "different kinds of *data*, and not of different kinds of *persons*."[111]

Raab and Bennett suggest that different sorts of social identities might be specified in the context of the kinds of relationships that characterize different sectors within the society. Thus there are groups that can be defined as customers, patients, suspects, students, and taxpayers. There are also groups that can be defined on the basis of still other categories that exist outside explicit relationships, but along some continuum: young/old, rich/poor, native/immigrant, etc. The complexities emerge, as we have noted, with regard to complex identities, as individuals navigate a great variety of sectors, and interact with institutional others from behind an assortment of presentational masks. As Raab and Bennett suggest, no simple aggregation can even represent the essence of a person.[112]

Surveys of consumers indicate a high level of concern about the loss of privacy. Yet, the numbers of respondents who say they have been victims of an "invasion of their privacy" is compara-

---

present banners with discounted offers to pre-selected browsers may be both able and willing to supply the lists of those who are given offers. The multitudes who are *not* provided the discount of special opportunity are unlikely to be supplied.

109. *See* Raab & Bennett, *supra* note 20, at 274.

110. *Id.* at 265.

111. *Id.* at 266.

112. *See id.* at 267.

tively low.[113] The surveys reflect a high level of perceived risk. Perceived risk is the basis upon which individuals decide whether to venture into transactions with strangers, or even decide to invest in the technology that would enable such transactions. Perceived risk may explain some part of the persistence of a "digital divide." This risk may be distributed unequally. However, when we engage the problem of risk, we are necessarily taken out of the realm of objective statistics, and into the uncertainty of perception. While there is still debate about the nature and extent of racial discrimination, there are at least standards, tests, and measures upon which comparisons can be made. There are no such standards in the area of perceived risk. The perception of risk is shaped by a complex of factors including the nature of the hazard and the way it has been portrayed in the press,[114] as well as by the sense of power and efficacy that varies systematically with race, gender, age and income.[115] This complexity ensures that no broad standards will ever appear.

In the context of discrimination more generally, it has been suggested that special protections for social groups and their members should be based on their actual or potential political power. The assumption here is that groups with political power can use that power to pursue their interests in the courts or through legislative representatives. Kenji Yoshimo elaborates on this argument in the context of the government's "don't ask, don't tell" policy governing homosexuals in the military.[116] After suggesting that the courts might withhold the grant of "heightened scrutiny" from groups that can change or conceal their defining traits, Yoshimo argues that neither mutability nor invisibility are really determinants of political power.[117] It is only

113. A survey administered mid-year in 1998 asked consumers "how concerned are you about threats to your personal privacy in America today." Some 87% said they were concerned, and more than half said they were very concerned. Yet, when asked if they had ever personally been the victim of "an improper invasion of privacy by a business" only 41 % said that this had happened to them. *See* Center for Social and Legal Research, *P&AB Survey Overview: Consensual Marketing is Coming*, 6 PRIVACY & AM. BUS. 1 (1999).

114. *See* ELEANOR SINGER & PHYLLIS ENDRENY, REPORTING ON RISK: HOW THE MASS MEDIA PORTRAY ACCIDENTS, DISEASES, DISASTERS, AND OTHER HAZARDS (1993).

115. *See* James Flynn et al., *Gender, Race, and Perceptions of Environmental Health Risks*, 14 RISK ANALYSIS 1101, 1108 (1994); *see also* Roger E. Kasperson & Jeanne X. Kasperson, *The Social Amplification and Attenuation of Risk*, 545 ANNALS, AAPSS 95, 105 (1996).

116. Kenji Yoshimo, *Assimilationist Bias in Equal Protection: The Visibility Presumption and the Case of "Don't Ask, Don't Tell"*, 108 YALE L.J. 485, 571 (1998).

117. People who are subject to excessive attention by police, especially customs agents because of the way they look, are advised to change their

when invisibility is compelled, that the outcome is disempowerment.[118]

Yoshimo's ultimate recommendation is that the jurisprudence of heightened scrutiny should be reconstructed solely on the basis of an analysis of political powerlessness. Yet, a further distinction is made. The traditional view is characterized as being class based, in which women, blacks, and gays are members of disempowered classes. An emergent view is one that is based on classification (or identification). The distinction would not be based on classes, but on the classifications or distinctions that create the classes (race, sex, or sexual orientation). The expressed value, according to Yoshimo is one of symmetry: "[I]f blacks get heightened scrutiny, so must whites."[119] This is the view that appears to be in ascendance within the U.S. courts.

I recognize that where political power may be the appropriate standard with regard to the courts, it is obvious that a very different standard of power operates within the marketplace. Unless minority tastes are combined with considerable wealth or disposable income, they are not honored in the market.[120] That they are not heard in the market for information is even more troubling, for many of the reasons we have already reviewed.[121] The challenge is to transform the lack of political power that may justify heightened scrutiny within the courts into a comparable level of scrutiny within the marketplace.

VIII. TOWARD GENERAL STANDARDS

As scholars and activists work toward the development of the means for defining and measuring the actual, perceived, and dynamic vulnerability of groups within markets and public spheres, there remains a pressing need for us to establish a more general policy to govern transactions in cyberspace.

---

appearance. This is not empowering, but submission to the use of an invidious distinction.

118. *See id.* at 557.

119. *Id.* at 563

120. Substantial evidence has been developed that demonstrates the low value that advertisers place on gaining access to African American audiences. Radio stations that program to these audiences receive a smaller share of advertiser revenue for each rating point than do stations that program for mainstream audiences. *See* CIVIL RIGHTS FORUM ON COMMUNICATIONS POLICY, WHEN BEING NO. 1 IS NOT ENOUGH: THE IMPACT OF ADVERTISING PRACTICES ON MINORITY-OWNED & MINORITY-FORMATTED BROADCAST STATIONS (1999).

121. *See* Edward S. Herman, *Commodification of Culture, in* TRIUMPH OF THE MARKET 3, 11 (1995).

The Federal Trade Commission has settled rather selectively on a subset of the OECD privacy guidelines. As a result, they have weakened their usefulness as a standard. The Commission must be challenged to explain why they have pulled back so far from a standard that seems on its face to have been very well considered. Joel Reidenberg has argued persuasively for the implementation of these standards in the context of a withering critique of information practices within the direct marketing industry.[122] Reidenberg has also been clear in his objections to reliance on the marketplace to develop and enforce these standards. His works have undoubtedly fallen upon deaf ears. Perhaps the Commission might respond more favorably to the recommendations of social scientists, rather than legal scholars. Gary Marx has proposed twenty-nine questions that he suggests that those who engage in surveillance might ask themselves as they plan the collection and use of personal information.[123] The Federal Trade Commission, as well as the Federal Communications Commission, and perhaps the Office of Management and Budget should consider the adoption of questions of this sort as the basis of evaluating the information practices of those organizations for whom they have oversight responsibility.

122. *See* Joel R. Reidenberg, *Setting Standards for Fair Information Practice in the U.S. Private Sector*, 80 IOWA L. REV. 497, 516-23 (1995). He has updated his critique more recently in Reidenberg, *supra* note 99, at 792.

123. Gary T. Marx, *Ethics for the New Surveillance*, 14 INFO. SOC'Y 171, 174 (1998).