

# Thinking Outside the Box: Considering Transparency, Anonymity, and Pseudonymity as Overall Solutions to the Problems of Information Privacy in the Internet Society

TAL Z. ZARSKY\*

Preamble .....	1302
I. Transparency .....	1305
I.1. What Is Transparency? .....	1305
Required Opacity .....	1311
The Information “Hubs” – Location, Interaction, and Transaction ..	1313
I.2. Transparency and the Practical Harms to Society .....	1315
(a) Fear of Abuse and Criminal Misuse .....	1316
(b) Fear of Errors in Databases .....	1319
(c) Price Discrimination .....	1320
(d) Manipulation and the “Autonomy Trap” .....	1325
I.3. Transparency as “Leveling the Playing Field”? .....	1331
Leveling vs. Mining .....	1332
II. Anonymity .....	1334
II.1. What Is Anonymity? .....	1334
II.2. The Troubles of the Anonymous Society .....	1337
Anonymity’s Enemies .....	1339
III. Pseudonymity .....	1340
III.1. What is Pseudonymity? .....	1340
Traceable vs. Untraceable Pseudonymity .....	1341
Traceable Pseudonymity: How Can It Be Done? .....	1342
Pseudonymity and Identity .....	1344
III.2. Pseudonymity, the Data Flow, and the Detriments of Surveillance .....	1345
III.3. Pseudonymity and the Troubles of Anonymity .....	1349
III.4. The Troubles of Pseudonymity .....	1350
(a) Keeping Up the Walls .....	1350
(b) “Virtual Babies and Virtual Bathwater” .....	1353
Conclusion .....	1354

*“Connect nine dots within a square array, with four straight lines without lifting your pen from the paper, nor crossing the lines more than once. Clue: Think outside the box.”*

*— A famous brainteaser*

---

\* Resident Fellow, Information Society Project, Yale Law School. J.S.D., LL.M, Columbia Law School. LL.B/B.A in Law and Psychology (*summa cum laude*), Hebrew University of Jerusalem. This Essay is the final chapter of a Doctorate Thesis at Columbia Law School. I thank Eben Moglen, Lance Liebman, Michael Dorf, and Daniel Solove, as well as Jack Balkin and the ISP fellows for their assistance and insightful remarks. I also thank Dr. David Brin for his helpful critique of earlier drafts.

## PREAMBLE

Recent technological developments are changing our world and society in many ways. This is especially true with regard to the flow of information that pertains to specific individuals. Such personal information is now easily collected and analyzed through the use of sophisticated means and implemented in a variety of ways.<sup>1</sup> The enhanced flow of personal information creates a variety of problems that in turn cause public concern and even outrage. The world is changing fast, and fears of Big Brother and Little Brother<sup>2</sup> are rampant again. Headlines announce “The Death of Privacy!” while the public tries to come to terms with Internet cookies, surveillance cameras, E-Z Pass records, supermarket cards and the prospect of biometric identifiers. This new reality of constant surveillance causes several layers of privacy concerns; not only does it lead to concerns stemming from the anxiety that constant surveillance creates, it also leads to fears of the actual uses of such personal data — such as fears of abuse or misuse, and even suspicions of discrimination and manipulation.<sup>3</sup>

In recent years, legislators,<sup>4</sup> courts, and legal scholars began to address these problems at length and to offer solutions.<sup>5</sup> Many of these solutions address privacy concerns by inhibiting the collection of spe-

---

1. For a discussion about the flow of personal information through its collection, analysis, and implementation, see Tal Zarsky, *Desperately Seeking Solutions — Using Implementation-Based Solutions for the Troubles of Information Privacy in the Age of Data Mining and the Internet Society*, 56 ME. L. REV. 13 (2004).

2. “Little Brother” is a phrase used by some to refer to surveillance carried out by private entities, as opposed to “Big Brother,” which connotes state surveillance.

3. See Tal Zarsky, “*Mine Your Own Business!*”: *Making the Case for the Implications of the Mining of Personal Information in the Forum of Public Opinion*, 5 YALE J.L. & TECH. 4, 17-47 (2003) for a taxonomy of current problems arising with information privacy.

4. Some U.S. laws addressing the collection of personal information are the Fair Credit Reporting Act of 1970, Pub. L. No. 90-32, 15 U.S.C. § 1681; the Cable Communications Policy Act of 1984, Pub. L. No. 98-549, 47 U.S.C. § 551; the Video Privacy Protection Act of 1988, Pub. L. No. 100-618, 18 U.S.C. §§ 2710-2711; the Driver’s Privacy Protection Act of 1994, Pub. L. No. 103-322, 18 U.S.C. §§ 2721-2725; the Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191; the Children’s Online Privacy Protection Act (COPPA) of 1998, Pub. L. No. 106-170, 15 U.S.C. §§ 6501-06; and the Gramm-Leach-Bliley Act of 1999, Pub. L. No. 106-102, 15 U.S.C. §§ 6801-6809. For a description of privacy-related laws outside the United States, see Mike Hatch, *The Privatization of Big Brother: Protecting Sensitive Personal Information from Commercial Interests in the 21st Century*, 27 WM. MITCHELL L. REV. 1457, 1468-71 (2001).

5. There are a great variety of writings on the problems and solutions to information privacy matters in today’s world. In general, see DANIEL J. SOLOVE & MARC ROTENBERG, *INFORMATION PRIVACY LAW* 1-62 (2003), and Lawrence Lessig, *Cyberspace and Privacy: A New Legal Paradigm?*, 52 STAN. L. REV. 987 (2000).

cific forms of data,<sup>6</sup> restricting its analysis,<sup>7</sup> or forbidding specific uses of the results of such analysis.<sup>8</sup> In other words, these forms of solutions are best understood as set out along the flow of personal information — from the time of its collection, the extent of its analysis, and permitted forms of its use. Obviously, interfering with or changing any one of these stages prevents the final problematic outcomes and is therefore a plausible solution. As I explain elsewhere,<sup>9</sup> presenting privacy concerns and solutions as part of the “information flow paradigm” — which consists of collection, analysis and use of the data — is extremely useful for understanding the implications of every privacy solution as well as the relation between the regulatory schemes.

Beyond these “standard” answers, recent research indicates other paradigms for solutions that involve “thinking outside the dataflow box” — *transparency*, *anonymity*, and *pseudonymity*. These solutions cannot be classified as part of the conventional “information flow paradigms” for solutions (as they do not directly restrict the collection, analysis or use of personal information), but in their essence present an alternative flow of information or even require the construction of an alternative society. Nevertheless, they provide an interesting response to the ills of our new information environment. They also offer many advantages in comparison to solutions situated inside the box, as they need not confront several of the legal contests the conventional solutions must face.

This Essay provides a novel attempt to address all three of these global solutions as a whole. In the following sections, I analyze various aspects of transparency, anonymity, and pseudonymity and envision societies where our personal information is, respectively, out in the open, hidden in its entirety, or somewhat blurred due to the use of multiple identities. I then address the way every one of these theoretical solutions attempts to resolve several privacy concerns arising in today’s

---

6. For example, the Cable Communication Policy Act (1994) prohibits the collection of information that is not necessary to maintain the cable service. 47 U.S.C. §551(c). In addition, collection restrictions are part of COPPA as well. U.S.C. §§ 6501-06.

7. For example, see the European Union (EU) regulation *infra* note 55.

8. For example, recent laws have restricted the practices of mass mailing and telemarketing — practices largely facilitated by the collection and analysis of personal information. *See, e.g.*, the Telephone Consumer Protection Act of 1991, Pub. L. No. 102-243, 47 U.S.C. § 227; Telemarketing and Consumer Fraud and Abuse Prevention Act, Pub. L. No. 103-297, 15 U.S.C. § 6101; and the Federal Trade Commission’s (FTC) telemarketing sales rules, 16 C.F.R. § 310.4. Most recently, the FTC and the Federal Communications Commission (FCC) have established a “Do Not Call” list and set high fines for commercial entities that contact individuals who specifically indicated they do not wish to receive solicitations. For a more complete view of the regulatory framework of the “Do Not Call” list, see <http://www.ftc.gov/bcp/rulemaking/tsr/tsrrulemaking/index.htm>. For a discussion of the legal challenges to this regulatory framework, see [http://www.reclaimdemocracy.org/corporate\\_speech/ftc\\_call\\_list\\_legal\\_analysis.html](http://www.reclaimdemocracy.org/corporate_speech/ftc_call_list_legal_analysis.html).

9. See Zarsky, *supra* note 1, at 17.

world of ongoing collection of personal information. The Essay starts by discussing and analyzing transparency (Part I), but finds it unsuitable as a response to today's privacy concerns, because the provision of vast amounts of personal information comes with incurable side effects. From there I move on to anonymity (Part II), which solves all problems but brings the flow of information to a halt and creates powerful side effects of its own. Finally, I settle for an intermediate solution — pseudonymity (Part III) — that solves many of today's privacy concerns and even empowers individuals in a unique and surprising way. Furthermore, pseudonymity offers several layers of privacy protection that can be partially removed when foul play is suspected, thus allowing surveillance to take place while some privacy is maintained. Carrying out an analysis of these global solutions in a single essay allows for an easy comparison among these extreme perspectives and their benefits and detriments, while using the same terms, models, and analytical structures for all three solutions. Moreover, by simultaneously understanding the benefits and detriments of the extremes — the lack of information or its overall disclosure — we will achieve a better understanding of the tensions that today's information flows create.

Beyond the obvious technological innovations currently facilitating the flow of personal information, my analysis focuses on two technological and social developments — the emerging Internet society, and the introduction of data mining applications that allow for a sophistication leap in the ability to carry out information analysis. The Internet provides greater opportunities for the collection of personal information about the habits and traits of individuals.<sup>10</sup> As a digital medium, it also facilitates the storage, warehousing, and analysis of such personal information.<sup>11</sup> Furthermore, various traits of the Internet allow for effective uses of personal information, in ways that are at times detrimental to the end user.<sup>12</sup> The Internet also empowers society by promoting accessibility to information resources. Therefore, much of the following analysis is set in the Internet realm — while addressing how transparency-, anonymity-, and pseudonymity-based schemes could be facilitated by the Internet infrastructure, and whether these solutions can successfully confront the unique privacy concerns the Internet environment creates.

---

10. For explanations and demonstrations as to how the shift to the online realm directly exacerbates privacy concerns, see Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1198 (1998), and Paul M. Schwartz, *Privacy and Democracy in Cyberspace*, 52 VAND. L. REV. 1609, 1628 (1999).

11. For a fuller explanation of the importance of efficient storage and warehousing, see Zarsky, *supra* note 3, at 8-9. See also BHAVANI THURASINGHAM, *DATA MINING — TECHNOLOGIES, TECHNIQUES, TOOLS, AND TRENDS* 49 (1999).

12. See analysis in Zarsky, *supra* note 3, pt. II. Also, see analysis below about the detrimental effects of online discrimination and manipulation.

Data mining is the enhanced ability to analyze vast amounts of personal information. The data mining tools, also known as Knowledge Discovery in Databases, or KDD applications, search for patterns and clusters within datasets, without receiving a hypothesis from the analyst and do so in an almost fully automated process.<sup>13</sup> The availability of data mining tools changes the information landscape, as it allows the user to derive more knowledge from less data. In this Essay, I directly examine how the availability of these tools affects the feasibility of the “outside the box” solutions.

The analysis focuses on the interaction between individuals and commercial entities. However, it also closely examines the government’s attitude toward these global solutions. The contemplated shifts toward these global solutions will be very difficult to achieve if strongly opposed by government, and to a greater extent depend on governmental endorsement for their successful implementation. Among other issues, these proposed solutions have serious implications for law enforcement and counter-terrorism activities, and therefore pose serious questions to governments and societies — questions that must be fully considered before these solutions’ enactment.

In all, this Essay provides a journey through the uncharted waters of problems and solutions in the world of personal data flows. Let us begin our journey by first examining the transparent society.

## I. TRANSPARENCY

### I.1. *What Is Transparency?*

A transparent society will feature constant and broad surveillance, but provide everyone with access to the outputs of such surveillance. In this society, privacy regulation will not limit collectors in their collection efforts, but require them to share the data they gather, as well as provide the public with additional information.<sup>14</sup> The conceptual cornerstone of transparency-based privacy solutions is that, in such regulatory schemes, the general public will be provided with equal access to the vast databases of personal information currently constructed by governments and commercial entities. In other words, the data streaming in from surveillance cameras and supermarket club cards should be made

---

13. For a layman’s version of the functions and use of some data mining applications, see Zarsky, *supra* note 3. For a definition of data mining and its basic applications, see U. M. FAYYAD ET AL., FROM DATA MINING TO KNOWLEDGE DISCOVERY: AN OVERVIEW, *in* ADVANCES IN KNOWLEDGE DISCOVERY AND DATA MINING 6 (Usama M. Fayyad et al. eds., 1996); *see also* DAVID HAND ET AL., PRINCIPLES OF DATA MINING (2001).

14. See DAVID BRIN, THE TRANSPARENT SOCIETY: WILL TECHNOLOGY FORCE US TO CHOOSE BETWEEN PRIVACY AND FREEDOM 80-84 (1998) for the concept of “Reciprocal Transparency.”

accessible not only to collectors, but to those subject to collection as well. As I explain below, the notion of a transparent society could be easily understood when shifting our perspective to the Internet realm, where vast amounts of personal data are collected, but potentially shared with ease. The transparency solution does not fit within the information flow paradigm (namely, addressing privacy concerns by regulating the collection, analysis, and use of such data), since it does not impede the data flow, but drastically enhances it.

The concept of the transparent society as an overall cure to today's personal information ailments is to a great extent the brainchild of David Brin, an accomplished science fiction writer and scientist.<sup>15</sup> In the *Transparent Society*, Brin advocates this idea in full force. He starts by discussing the current practices of surveillance through the spreading use of cameras and other means of data collection.<sup>16</sup> To solve the problems such collection creates, Brin suggests that rather than closing down information flows by inhibiting the collection of data in various circumstances, access to such data should be provided to all — thus enhancing the flow of personal information.<sup>17</sup> For example, instead of restricting the rights of large companies to collect data about their consumers, such companies should be allowed to do so on the condition that the company's top hundred officers post exactly the same form of information currently collected, as it pertains to them and their families, on the Internet. (This example is indeed radical and even irrational, but it provides some direction for understanding the possibilities of transparency-based regulations.) He goes on to suggest that rather than securing online transactions by encryption and e-cash, such transactions should be carried out in the open.<sup>18</sup> Brin's rationale for such radical suggestions is that the existence of transparency will deter foul play and abuses of the use of personal data, as all actions will be viewable, detectable, and above all, traceable.

With transparency in place, Brin believes that society will protect itself from the detriments of the enhanced information flow.<sup>19</sup> For example, neighborhood watches will overlook the city streets through the data flowing online from centrally situated cameras.<sup>20</sup> Moreover, transparency will create a balance of power between all parts of society that are now both collectors and subject to collection, and such balance

---

15. See generally *id.* For more about David Brin, see <http://www.davidbrin.com>.

16. See BRIN, *supra* note 14 at 5-8.

17. *Id.* at 81.

18. *Id.* at 182-84.

19. *Id.* at 258-60.

20. *Id.* at 258.

will minimize conflicts.<sup>21</sup> Brin states that in a transparent society, individuals will retain a sense of participation and control, rather than the distrust and fatalism that arise when data collection and use are covert.<sup>22</sup> Naturally, Brin addresses the most apparent fear stemming from a transparent society — the fear that people will constantly and obsessively watch over others if they have the means of doing so. But he concludes that, in a society of full disclosure, people will learn to ignore the vast amount of personal information made available and find better uses for their time than spying on their peers.<sup>23</sup> Since everyone has something to hide, no one will look for the flaws of others.

It should be noted that referring to transparency and disclosure as means of achieving social objectives did not begin or end with Brin's ideas. Several laws and regulations already in place rely on mandatory disclosure to promote equality, enhance performances, generate efficiency, and assure a fair democratic process.<sup>24</sup> The most salient example is the Freedom of Information Act (FOIA),<sup>25</sup> which provides citizens with a right to monitor the actions of government. However, the FOIA is a weak example of transparency, at least in the robust context currently addressed, as it includes a personal privacy exemption. According to Exemption 6 of the Act,<sup>26</sup> when a request is made for records containing information about a private citizen, nondisclosure is mandated if the invasion would be clearly unwarranted. Therefore, the courts are required to weigh the personal privacy interest at stake against the public interest in disclosure.<sup>27</sup> Other laws and regulations premised on transparency and disclosure<sup>28</sup> are rules of financial disclosure, Toxic Reduction Inventory rules, nutritional labeling, the Home Mortgaging Disclosure Act, and the Federal Election Act.

However, the transparent society Brin envisions goes far beyond these rules of disclosure, both in its scope and ambitions. It does not limit itself to specific sectors, but reaches across the entire society. It brings the spotlight of public scrutiny to the actions of every one of us, not only to large corporations or public figures. Initially, this idea was

---

21. *Id.* at 254. Brin's analogy in this context is to the balance of power between the two superpowers during the Cold War, which eventually led to a period of relative peace.

22. *Id.* at 154.

23. *Id.* at 298.

24. On this issue, see David Weil, *The Benefits and Costs of Transparency: A Model of Disclosure Based Regulation* (June 2002), available at <http://ssrn.com/abstract=316145>, Table 1.

25. 5 U.S.C. § 552 (1998).

26. *Id.* § 552(b)(6).

27. See generally Glenn Dickinson, Comment, *The Supreme Court's Narrow Reading of the Public Interest Served by the Freedom of Information Act*, 59 U. CIN. L. REV. 191, 196 (1990) for a discussion as to the Court's rulings and attitude on these issues.

28. See Weil, *supra* note 24.

viewed as somewhat of a science fiction concept, but it has been gaining popularity and recognition as a serious solution to the problems stemming from the collection and use of personal information in today's society.<sup>29</sup> Obviously, this transparent reality would be very different from the present situation, which resembles a one-way mirror where only a part of society is eagerly collecting and using personal information, while the silent majority remains mostly passive and uninformed.

The underlying concepts leading Brin and others to prefer the transparency-based solution over other forms of privacy-enhancing solutions (including those mentioned above as being "inside the box") are those of equality and accountability. Transparency strives to achieve equality by leveling the playing field between individuals as well as between the segments and classes of society and providing them equal access to the personal information of others. With transparency in place, information asymmetries will be minimized, thus creating efficient and fair markets.<sup>30</sup> Furthermore, the underlying assumption of those advocating transparency is that equality in the distribution of and access to information will lead to equality in the distribution of power and control in society.<sup>31</sup> For example, in a transparent society, large businesses and small consumers are on somewhat even ground, as both can access and use the same information as well as apply data analysis and data mining tools to the information they obtain. While businesses can use the personal information the transparent society provides to make educated guesses and predictions about the future conduct of their customers, consumers in response can seek out information about various firms, searching for their weaknesses and vulnerabilities. Thus, both sides will be equally equipped for participation in the market.

Another key concept of transparency is accountability. In a transparent society, all actions are monitored, and anyone carrying out abusive or malicious acts will be held accountable for his or her actions, thus mitigating any abusive uses of the vast amounts of personal information now available. As I explain below, Brin relies on this element

---

29. Even though somewhat subjective, see David Brin's website at <http://www.davidbrin.com/privacyarticles.html> for a list of follow-up interviews, discussions, and related articles.

30. Either symmetric information or the lack of information promotes an efficient functioning of the market. However, asymmetric information causes market deterioration. For a further analysis of these points with regard to insurance and financial markets, see Eric K. Clemons et al., *Impacts of e-Commerce and Enhanced Information Endowments on Financial Services: A Quantitative Analysis of Transparency, Differential Pricing and Disintermediation*, 22 J. FIN. SERVS. RES. 73 (2002).

31. As Francis Bacon famously said: "*scientia est potentia*" — knowledge is power. By addressing inequity in access and distribution of knowledge, inequities in the distribution of power should be affected as well.

heavily, and not always correctly.<sup>32</sup> Accountability and transparency are often mentioned with regard to oversight of state actions,<sup>33</sup> as it is generally assumed that with transparency in place, the state will refrain from abusing and misusing its powers, because it will be held accountable for its actions.<sup>34</sup> In this Essay, however, I look beyond these government-related issues, and emphasize the problems arising from the collection and analysis of personal information by private parties.

In theory, the transparency solution appears promising and even seductive in its appeal. It is promoted as the silver bullet that will potentially mitigate many of today's privacy concerns. It allows society to enjoy the benefits of information collection and analysis<sup>35</sup> while remaining protected from possible abuses. Moreover, the transparency scheme could be implemented without elaborate regulations and expensive enforcement. There is, however, a large speed bump on the road to transparency's success: It will take a great deal of convincing to persuade legislators, academia, and society in general to accept this form of solution since the transparency paradigm presents a reality that is the almost extreme opposite to today's legal and social understanding of privacy. Currently, the dominant trends of thought with regard to privacy focus on the troubles arising from collection *per se* of personal information. Privacy scholars holding these views emphasize the detriments of constant monitoring on its own, and how it affects the individual's state of mind, inhibits daily activities,<sup>36</sup> promotes conformity, causes embarrassment, and interferes with the creation of intimacy.<sup>37</sup> Others stress the importance of preserving unmonitored choices to assure the individual's autonomy.<sup>38</sup> But almost all of those who assert that these perspectives point to the collection of information as the problem and the harm itself (and not only a potential for harm that such

---

32. I discuss below whether such accountability is indeed effective as a shield against the possible ills arising in a transparent society.

33. See Dickinson, *supra* note 23.

34. See BRIN, *supra* note 14, at 108-10.

35. See Zarsky, *supra* note 1, pt. II, for an extended discussion as to the benefits that can derive from the analysis of personal information. Such benefits include subsidies to startups and additional value to both vendors and consumers from the analysis of personal information.

36. See, e.g., *id.*, pt. II; Daniel Solove, Conceptualizing Privacy, 90 CALIF. L. REV. 1087, 1130 (2002).

37. See, e.g., Kang, *supra* note 10. Kang argues that the lack of privacy in personal information denies the individual the ability to trickle out personal information and in that way to invite and affirm intimacy. Kang suggests a critique to this claim as well — that the creation of trust and intimacy is related to the sharing of experiences, rather than of secrets (based on Jeffrey H. Reiman, *Privacy, Intimacy, and Personhood*, in PHILOSOPHICAL DIMENSIONS OF PRIVACY: AN ANTHOLOGY 300 (Ferdinand David Schoeman ed., 1984)).

38. See, e.g., Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1425 (2000).

collection of data can cause) will surely disagree that accelerated collection and sharing of persona data will lead to any solution. In addition, it will also be difficult to convince the public<sup>39</sup> to embrace a solution that promotes transparency and requires society to ignore the seemingly innate human trait of seclusion.<sup>40</sup> In view of these opinions, it is clear why discussions about the use of transparency as an overall solution are rare, and the idea itself is considered more of a science-fiction concept than a viable legal and social doctrine.

Regardless of these powerful counterarguments, the current resentment to constant surveillance, and the gloomy prospects for success, transparency can prevail. With new technologies emerging, society is capable of adapting and changing its expectation of privacy, as it has done in the past.<sup>41</sup> What is unclear, however, is whether transparency can achieve the objectives it set out to meet. I therefore devote the following pages to examining whether a wide use of transparency solves, or perhaps accidentally intensifies, the actual problems that the collection and analysis of personal information brings about through the creation of accountability, equality, and other side effects of this solution. I also examine whether equality in the access to information will lead to equality in power. As previously mentioned, I focus the following analysis on the somewhat narrow Internet context. I conclude that even though *scientia* (knowledge) leads to *potentia* (power), there is still a significant gap between data and insightful knowledge — a gap that is widened by the emergence of data mining tools.

To facilitate the analysis, I begin by creating an overall picture of what a transparent society might realistically entail in accordance to Brin's suggestions and of which forms of information it will provide to all, given today's social and legal constraints. To describe this hypothetical society, I borrow several of Brin's paradigms, filling in voids where

---

39. For insights on current public opinion on these issues, see *Opinion Surveys: What Consumers Have to Say About Information Privacy: Hearing Before the House Comm. on Energy and Commerce and the Subcomm. on Commerce, Trade and Consumer Protection*, 107th Congress (2001) (statement of Dr. Alan Westin), available at <http://energycommerce.house.gov/107/hearings/05082001Hearing209/Westin309.htm>. In general, and according to Westin, such surveys indicate that the public is indeed concerned with privacy — especially with regard to intrusion (unwanted mail and telemarketing), manipulation (profiling that allows “hidden persuader” marketing), and discrimination.

40. Note however that the privacy we have grown accustomed to today is very different from the expectations of previous generations, especially from that of Colonial times in America. See, e.g., ROBERT ELLIS SMITH, BEN FRANKLIN'S WEB SITE 18 (2000). In addition, note that the standards of privacy vary even today in different parts of the world.

41. This opinion is advocated by ESTHER DYSON, RELEASE 2.1: A DESIGN FOR LIVING IN THE DIGITAL AGE 274 (1998). Dyson opines that people will grow accustomed to the fact that they are reviewed and recorded on the Internet, eventually resulting in a new form of privacy expectancy.

required, and at times broadening or narrowing the realms of privacy and transparency.

#### REQUIRED OPACITY

Calling for a society that is transparent in its entirety is clearly an exaggeration, as information flows should clearly cease at specific times and places.<sup>42</sup> In any transparent society, a realm of privacy and seclusion will still persist around individuals, and remain with regard to specific forms of actions, transactions, and communications.<sup>43</sup> Drawing out this theoretical realm of seclusion is a challenging task. Below, I try to predict the nature and extent of these realms:

(a) On the most basic level, the public's vision should not penetrate the individual's home, and transparency requirements should not oblige individuals to disclose information concerning their health or family. The concept of the home and family as safe havens from the prowling eyes of others is deeply embedded in social and legal norms,<sup>44</sup> and any new global solution will surely prove unsuccessful in changing or challenging this conception. Moreover, the establishment of basic forms of trust among family members requires that these realms of seclusion remain intact, regardless of any transparency scheme.<sup>45</sup> As the social norms regarding the Internet are constantly in flux, it is difficult to establish what will consist of the "home" or "a zone beyond collection" in this context. Future research and writing will surely address how new forms of privacy and identity are formed in the online environment.

(b) A transparent society will be sure to include privileged relations, which will restrain trusted parties from providing information about specific people and issues.<sup>46</sup> In addition to the privileges that are widely accepted today, such as lawyer-client and physician-patient, a broad transparency scheme will require privilege-like protection for additional relationships such as bank-client<sup>47</sup> or employer-employee.

---

42. For one perspective on these zones of opacity, see Richard S. Murphy, *Property Rights in Personal Information: An Economic Defense of Privacy*, 84 GEO. L.J. 2381, 2409 (1996).

43. See BRIN, *supra* note 14, at 210.

44. The protection of the individual's home from the prowling eyes of the government has been recently revisited in *Kyllo v. United States*, 533 U.S. 27 (2001). The Supreme Court held that the use of thermal imaging heat detectors by the police to view the interiors of private homes in order to detect marijuana-growing lights was unconstitutional. The privacy right within the individual's home is also protected from non-governmental entities. See, e.g., RESTATEMENT (SECOND) OF TORTS: INCLUSION UPON SECLUSION § 652B. See also Zarsky, *supra* note 1.

45. For a discussion of privacy and intimacy, see *supra* note 37.

46. On these issues, see discussion in Zarsky, *supra* note 1, pt. IV.

47. The scope of this privilege is to a certain extent addressed in the Bank Secrecy Act of 1970, Pub. L. No. 91-508. On this issue, see DANIEL J. SOLOVE & MARC ROTENBERG, *INFORMATION PRIVACY LAW* 525 (2003).

Such privileges should not, however, be extended to include information flowing from the relations between buyers and sellers, or other business interactions, since such privileges would block most of the information that is required for the full effects of transparency.<sup>48</sup>

(c) The contents of communications between parties — actual conversations, phone calls, letters or emails — should remain protected and outside the public domain even with a shift to overall transparency. Therefore, third parties — those facilitating the transfer of information and communications between individuals — should be prohibited from tapping into or publicizing the actual content of such communications. To a certain extent, defining the boundaries of the transparent society requires us to revisit the distinction between secrecy and anonymity.<sup>49</sup> Transparency will directly affect the user's ability to remain anonymous, as the occurrence of the conversation or transmission will become common knowledge.<sup>50</sup> However, transparency should still allow full secrecy in the content of communications.

(d) A transparent society should allow corporations to conceal certain forms of commercial information in the interest of enabling a successful market and economy. This requirement stems from the economic rationale that maintains that to achieve an efficient market and efficient business practices, regulation should promote secrecy and privacy among businesses, and disclosure among customers and individuals.<sup>51</sup> Businesses require a minimal level of privacy to allow them to refrain from revealing their business practices, thus facilitating an innovative and competitive market. Therefore, any model of transparency should allow today's trade secrets doctrines<sup>52</sup> to persist, and should preserve opacity in minutes of internal corporate meetings and position papers written by company executives.

It should be noted that recent corporate scandals<sup>53</sup> have again

---

48. In the Internet context, proposed unique privileges will of course include the relation between users and their ISPs or broadband providers. I address this issue below in Part III with regard to intermediaries that facilitate pseudonymity online.

49. I return to this distinction when discussing the global solutions of anonymity and pseudonymity (Parts II & III below).

50. The distinction between intercepting ongoing conversations or stored information and gathering information as to the source and destination of an electronic message is also evident in the Electronic Communications Privacy Act (ECPA) of 1986, Pub. L. No. 99-508, 18 U.S.C. § 2510. Title I and II of the ECPA regulate interceptions and access to stored messages and present law enforcement with stringent requirements. Title III of ECPA regulates pen registers and track and trace devices — namely the collection of information about the sender and recipient of data, and presents government with more lenient requirements. For an in-depth presentation of this issue, see SOLOVE & ROTENBERG, *supra* note 43, at 324.

51. See generally Murphy, *supra* note 42, at 2382; see text accompanying notes 15-17.

52. See MARTIN J. ADELMAN ET AL., PATENT LAW 51-73 (1998).

53. For example, the recent Enron and WorldCom debacles.

proved the importance of broad and true disclosure in the corporate setting. However, this Essay does not discuss issues of corporate governance and disclosure, since they are addressed by specific securities and corporate legislation and are subject to other forms of pressure groups, policy constraints, and market forces.

(e) Finally, when constructing the zones of opacity that will remain outside transparency's realm, public policy will surely require that specific forms of information should remain secure and anonymous. For instance, such concerns will surely be voiced with regard to personal data pertaining to children, who are considered vulnerable and perhaps gullible.<sup>54</sup> Concerns and demands for anonymity will also arise regarding additional forms of behavior or speech, such as information concerning voting and elections, and specific personal factors such as race, sexual orientation, or membership in unions and other specific organizations.<sup>55</sup>

THE INFORMATION "HUBS" — LOCATION, INTERACTION,  
AND TRANSACTION

When taking into account these zones of opacity, it is clear that no transparent society will be as transparent as we might have originally imagined, but it will still potentially facilitate the sharing of vast amounts of personal information. To understand the extent of the information available in a transparency scheme as well as its possible uses, both beneficial and detrimental, I introduce a three-hub model, including information about the *locations, interactions, and transactions* of individuals.

*Location* is the first form of information that transparency will surely provide. In a transparent society, data concerning the whereabouts of individuals will be open to all. Such information will stream in from E-Z pass points, cameras equipped with face recognition software situated in public areas such as Times Square and in private areas that are accessible to the general public, such as malls or large department stores.<sup>56</sup> However, private establishments of medium to small size will not be required to disclose this form of information. The location data

---

54. The protection of personal information pertaining to children is an issue directly addressed by legislation. See Children's Online Privacy Protection Act, 15 U.S.C. §§ 6501-6506. This act introduces restrictions on the collection of such information. See generally SOLOVE & ROTENBERG, *supra* note 47 at 561-63.

55. In the EU, the European Directive of Data Protection, effective October 25, 1998 (Directive 95/46/EU). Art. 8(1) specifically notes several categories of information that may not be collected or analyzed — including union membership.

56. See Zarsky, *supra* note 1, at 17, for a full description of such means of collection in the offline and online worlds. For a discussion of the use of surveillance cameras, see, e.g., Jeffrey Rosen, *A Watchful State*, N.Y. TIMES MAGAZINE, Oct. 7, 2001, at 38.

will allow anyone interested to determine who is at a specific place, at any time. In the Internet context, this information will include data as to what websites and chat rooms users are visiting, thus providing an overall log of their location in virtual space.<sup>57</sup>

*Interaction* — the second form of personal information available in a transparent society — will detail which individuals interact with others. To some extent, this source of information is a derivative of the location data described above. An analysis of location data will allow the analyst to assume that an interaction has occurred when two or more individuals are at the same place, be it physical or virtual, or in close proximity, at the same time.<sup>58</sup> Interactions could also be deduced from communications between different parties, such as phone calls, letters, and emails. However, as mentioned above, transparency should not pertain to the *content* of the communications among the individuals, which will remain outside the public eye and the realm of such interaction information.<sup>59</sup>

Third, a specific and important form of “interaction” data will pertain to *transactions*. In a transparent society, vendors will provide public access to records detailing the transactions in which they participated, including lists of the parties whom they recently transacted with. At this time, however, I can only speculate as to what such disclosed records will contain and whether they will list the exact details of every transaction. There are a variety of options across the spectrum of disclosure, which I explore below.

In the not so distant past, it would have been very hard to imagine how data collectors could provide the general public with real-time, actual access to these information resources.<sup>60</sup> Today, however, we can easily envision how the Internet’s infrastructure can facilitate such information sharing — for data pertaining to both our offline and online lives. Live pictures streaming from cameras and access to newly formed databases could be provided online to Internet users, thus allowing everyone a view of the locations, interactions, and transactions of others. By these means, personal data is not only available but literally at the

---

57. For a lengthy description of the information collectable online through the use of “cookies,” see, e.g., *In re DoubleClick Inc. Privacy Litigation* 154 F. Supp. 2d 497 (S.D.N.Y. 2001).

58. Such analysis will also use corroborating data such as the length of the interaction and whether it has occurred and recurred several times.

59. See *supra* note 50.

60. Clearly transparency does not directly require the use or even existence of the Internet, as the relevant information could be introduced to the public domain with a variety of low-tech options (for example, posted in newspapers or companies’ offices). However such options do not provide realistic and actual accessibility of information to the entire public.

fingertips of many, thus potentially facilitating full transparency and leading to equality and accountability.

Before I analyze the effects of any hypothetical transparent society, it is important to note that regardless of governmental policy and regulation, the extent of personal data available in any transparent regime will be a result of the opinions, preferences, and morals of the specific individuals that make up the relevant society. Should a person have a great interest in keeping certain forms of information undisclosed, he or she will strive to do so even at the price of inconvenience (as in the use of cash) or additional expense.<sup>61</sup> Therefore, it is very hard to predict the actual disclosure that a transparent society will create, as the data provided will vary from one individual to another. For example, the realm of personal information concerning transactions that transparency will provide can range from full disclosure of the transaction to information merely stating that a transaction between two parties occurred — information that can even be derived from the interaction data. This latter, narrow option will result from the reluctance by some customers<sup>62</sup> and many vendors<sup>63</sup> to inform others of the actual contents of their purchase baskets as well as the actual price paid in any given transaction.

### I.2. *Transparency and the Practical Harms to Society*

Despite the constraints mentioned, the “diet” transparency that may emerge in a transparent society will provide a vast amount of information and advantages on the one hand, and varied problems on the other. Can this form of transparency provide a response to the information privacy concerns of today’s society? To properly analyze the advantages and problems of transparency, I first acknowledge which privacy concerns need to be addressed. As I mentioned above<sup>64</sup> and elsewhere<sup>65</sup> most scholars understand privacy concerns as those that stem from the existence of constant collection and surveillance. In this Essay, however, I choose to examine how transparency will confront a different set of privacy-based concerns — those stemming from fears of the actual detrimental uses of personal data collected by commercial entities. For

---

61. See BRIN, *supra* note 14, at 250, where Brin addresses this issue and suggests that those who wish to sustain the secrecy of specific issues and use encrypted messages should be taxed.

62. Note that in today’s legal and business environment, information about the content of consumers’ shopping baskets is constantly used by sellers, their affiliates, and even third parties (which purchase such data on the secondary market). However, I assume that customers will be more concerned with the availability of such information to their neighbors (as opposed to vague corporate entities), as these are the individuals they interact with daily.

63. See Clemons, *supra* note 30 (stating that price transparency is viewed as a serious threat to vendors as it can lead to increasingly brutal price competition).

64. See Part I.1 above.

65. See Zarsky *supra* note 1, at 32.

this matter, I make use of a taxonomy I discuss elsewhere,<sup>66</sup> which addresses the problems arising in today's world of enhanced collection of personal data. These problems include the following: (1) fear of abuse and criminal misuse of personal data; (2) the fear of errors in databases; (3) the use of personal information to discriminate between users; and (4) the use of such data by content providers and advertisers to manipulate and impinge on personal autonomy. Because the transparency solution is premised on providing additional personal data and simplifying its collection, these problems purportedly would worsen with full disclosure. However, proponents of the transparency solution will argue that the opposite is true and that transparency will mitigate and even eliminate such problems — in most cases by introducing society to an enhanced level of equality and accountability. In the following paragraphs, I examine these competing arguments in an attempt to reveal the effectiveness of a transparency solution to privacy concerns. I carry out such analysis while focusing on interactions and transactions carried out online, and I specifically address the availability of data mining applications to analyze personal information.

(A) FEAR OF ABUSE AND CRIMINAL MISUSE

Today's constant collection of personal data raises serious concerns that information now available may be used to embarrass, expose, blackmail, or facilitate other harms — problems I address in depth elsewhere.<sup>67</sup> Such fears could be directed both at the state and at private entities and individuals. These harms are reflected in a great deal of anecdotal evidence appearing in the privacy literature.<sup>68</sup> Transparency, as an overall solution to privacy concerns, must provide a suitable response to these concerns. However, at first blush, such concerns will naturally intensify in a transparent society. Transparency protagonists will argue, however, that the overall effects of transparency, in view of the equality and accountability it provides, may not only compensate for the additional information made public, but also substantially mitigate these problems.

The argument in favor of transparency in this context is as follows:

---

66. Zarsky, *supra* note 3, pt. II. In what follows, I address all the problems I note elsewhere as arising from the collection of personal information, with the exception of "intrusion on seclusion." *Id.* at 68. This problem is indeed evident in a transparent society, but is sufficiently addressed above in the analysis of Required Opacity (Part I.1).

67. See Zarsky, *supra* note 3, pt. II.

68. A salient example of such misuses is the MetroMail case, in which a woman received a harassing letter with intimate details about her life from an inmate she did not know. These details were available inside the prison, as the inmates were employed in transferring personal information from various sources to electronic databanks. See Nina Bernstein, *The Erosion of Privacy*, N.Y. TIMES, June 12, 1997, at A1.

With transparency in place, fears that the available personal information will be used to embarrass, expose, blackmail, or in general gain power over another are to a certain extent mitigated since this information is accessible by all. With all personal information exposed, the threat of publicizing such information becomes ineffective. Moreover, the fact that everyone can access information about everyone else with ease and efficiency (i.e., the equality element of transparency) will deter this form of misuse of personal information, as users will know they will be subjected to immediate retaliation, as their personal information is available and transparent as well. In his book, Brin compares this reality to an arms race — where everyone is armed with information about the other, a balance of power will emerge, leading to a stalemate, or a stop to all detrimental uses of personal data.<sup>69</sup> Another form of abuse that may generate concerns stems from the government's ability to tap into the vast amounts of personal information that transparency makes available, and possibly use this information for spying and harming many citizens. In this Essay, I do not address this important issue of transparency in the relation between state and citizen<sup>70</sup> directly, since I focus on private entities, but these concerns can be addressed through specific legislation pertaining to acts of government and should not necessarily undermine the transparency solution as a whole.

The argument that transparency will facilitate retaliation and thus deter abuse overlooks an important factor —not all of those subject to possible abuses are capable of retaliation. In addition, transparency will also assist criminals by allowing them to learn at what times persons are alone or vulnerable, and to take advantage of these situations. Therefore, in a transparent society drawn according to the lines described above, danger will await the young woman strolling alone in an abandoned but monitored alley, as well as the elderly gentleman with a tendency to make unwise business decisions, a fact that is evident from his public record of bad investments. Swindlers and other predators, both online and offline, will search for such vulnerable victims, presenting law enforcement agencies with serious challenges. In addition, transparency will facilitate identity theft — a growing problem in today's information environment.<sup>71</sup>

In his book, Brin acknowledges this problem, but claims that a

---

69. See BRIN, *supra* note 14, at 255.

70. Transparency is an important check on government, allowing citizens to examine the actions of government officials and to hold them accountable for those actions. See discussion above regarding the FOIA (and its limitations).

71. The severity of this problem motivated Congress to pass the Identity Theft and Assumption Deterrence Act of 1998, 18 U.S.C. § 1028 (2000). See generally Solove, *supra* note 47, at 515.

transparent society will offer a sufficient response in the form of *reciprocal protection*.<sup>72</sup> In Brin's transparent society, reciprocal protection will unfold, as no one will be alone and in a way we will all be part of law enforcement. We will all be watching each other's backs by means of open surveillance and data sharing, thus deterring criminal behavior or abuse. Such reciprocity will compensate for those who are unable to retaliate. We will also be watching the police and ensuring that they carry out their missions properly and fairly.<sup>73</sup> Clearly, this claim relies heavily on the accountability that transparency enhances.

Brin's argument, though interesting, is quite problematic. In reality, individuals do not rush to assist others in danger or need, especially when there is no clear indication that it is their duty to do so. Social psychologists refer to this phenomenon as the "bystander effect" and the "diffusion of responsibility."<sup>74</sup> Research has shown that when large groups of people witness acts of violence or anti-social behavior, responsibility tends to diffuse among the witnesses. The most famous example is the murder of Kitty Genovese, who was beaten for an extended period of time in New York City while several neighbors watched through their windows. None of these witnesses intervened or called the police. The common explanation is that the silent neighbors assumed that someone else was attending to this matter, thereby diffusing their responsibility.<sup>75</sup> Experiments have found that the more bystanders witness an act, the less chance that any one of them will step forward.<sup>76</sup> In the transparency schemes discussed above, especially those involving the sharing of information through the Internet where the whole world is potentially watching, the chances for one specific

---

72. BRIN, *supra*, note 14, at 80.

73. This point goes to the issue of transparency in the actions of government, and thus is beyond the scope of the current discussion.

74. These issues were introduced in JOHN DARLEY & BIBB LATANE, *THE UNRESPONSIVE BYSTANDER: WHY DOESN'T HE HELP?* (1970).

75. *Id.* See also J. Darley & B. Latane, *Bystander Intervention in Emergencies: Diffusion of Responsibility*, 8 J. PERSONALITY & SOC. PSYCHOL., 377, 383 (1968). In their groundbreaking research (motivated in part by the Genovese murder), Darley and Latane established that intervention should not be viewed on its own but as the final step in a complex sequence of interpretations (see *THE HANDBOOK OF SOCIAL PSYCHOLOGY VOL. II*, at 44, 156, 282, 295 (4th ed. 1998)). They explained that in the Genovese case, the neighbors were looking for evidence that the situation was indeed an emergency. As they were not provided with any reassurance from others, the witnesses remained passive. To understand this situation, Darley and Latane introduced the following paradigm for the decision-making process that unfolds in these situations: The individual must establish that (1) a need to act exists; (2) the individual has a personal responsibility to act; and (3) there is something he or she could do to help. These decisions must be made under pressure when the natural default is to remain passive.

76. However, research proved that people *do* take action to help others when they are alone and understand that they are the only ones to witness the event (so the responsibility does not diffuse). *Id.*

individual with no designated duty to get involved voluntarily are therefore quite slim.<sup>77</sup>

An additional flaw in the concept of reciprocal protection is its underlying assumption that society's concern will spread equally across vast areas of interest and will remain attentive to the problems of all individuals. However, the reality is that human attention tends to focus on the sensational and dramatic. The public might be on the lookout for acts of violence in alleys of the large city, but will not be as interested with rural locations — both in the virtual and physical sense — and thus will leave non-urbanites vulnerable.<sup>78</sup>

In summation, a transparent society creates vast potential for criminal and other abuse. Law enforcement will stand no chance in handling such threats, and the vulnerabilities that transparency exposes must be countered by a shift in social norms of responsibility. Sadly, there is no guarantee that such a shift will indeed occur. The accountability transparency introduces might sound convincing in theory, but given the human traits addressed above, it has almost no chance of succeeding in practice. Thus, the transparent society would arrive at a very high price to public safety — a price we might not be interested in bearing.

#### (B) FEAR OF ERRORS IN DATABASES

Errors in the collected information and databases<sup>79</sup> can cause unfair treatment of specific individuals when the personal information is later used.<sup>80</sup> At first glance, transparency again exacerbates this problem by creating many more databases that are prone to include errors. Further analysis, however, leads to the opposite conclusion.

---

77. See *Net Grief for Online "Suicide,"* BBC NEWS, Feb. 4, 2003, available at [news.bbc.co.uk/1/hi/technology/2724819.stm](http://news.bbc.co.uk/1/hi/technology/2724819.stm), for a story that somewhat resembles the Genovese incident but occurred entirely online. In this case, a man boasted online of his use of drugs and evidently died of an overdose. Other "netizens" failed to react in time to save him, while others even dared him to continue with his drug usage.

78. Note that other theories of social psychology explain that the tendency to remain passive in emergencies is a result of the individual's information overload. This phenomenon is especially effective in explaining behavior in large cities, where people choose to ignore specific forms of information. *Id.* (citing Stanley Milgram, *The Experience of Living in Cities* 167 SCIENCE 1461 (1970)).

79. I address these issues in length both in Zarsky, *supra* note 3, pt. II, as well as in Zarsky, *supra*, note 1, pt. IV. For additional discussion of this issue, see, e.g., LAWRENCE LESSIG, *CODE: AND OTHER LAWS OF CYBERSPACE* 154 (1999). Here, Lessig states that in today's society, the burden of innocence has shifted to the individual, who must prove the database is wrong. See also SIMON GARFINKEL, *DATABASE NATION: THE DEATH OF PRIVACY IN THE 21ST CENTURY* 25 (2000) (discussing common errors in credit bureau databases and their devastating effects).

80. In Zarsky, *supra* note 3, I address a concern about errors that do not arise from tainted data, but from inherent problems in the process of data analysis, which causes the users to be treated in a biased manner. The analysis of this concern is beyond the scope of this discussion, and is not affected by a shift to transparency.

The oft-suggested remedy to this tragedy of errors has been a right of access to the collected databases, so that the relevant persons could examine the collected data and demand corrections when necessary.<sup>81</sup> The transparency solution addresses this problem directly by providing such a right of access to all (again it is the equality element that comes to our rescue).<sup>82</sup> Since all users can view the databases, everyone will be aware of errors and take action to correct them. Therefore, for transparency to sufficiently address the fear of errors, a right of correction must supplement the right of access that transparency will provide. With such a right in place, it appears that transparency can sufficiently mitigate the fear of errors.

### (C) DISCRIMINATION

An additional concern stemming from today's practices of collection and analysis of personal information is that the personal information now made available will facilitate discrimination among consumers.<sup>83</sup> With today's technologies of collection, analysis, and marketing, advanced forms of discrimination are made possible, especially within the Internet realm — an issue I will herewith address in depth.<sup>84</sup> The collection and analysis of vast amounts of personal information pertaining to consumers' patterns of behavior, in conjunction with Internet websites' ability to provide consumers with a customized shopping environment, allows marketers and vendors to discriminate among individuals with great precision and minimal effort. The "market-for-one" interface that the Internet infrastructure facilitates can even lock customers into higher prices<sup>85</sup> and thus improves the chances of such discriminatory schemes.<sup>86</sup> Discrimination is especially effective in situations

---

81. A right of access to the data pertaining to every individual is one of the five Fair Information Practices that the FTC indicates as "widely accepted principles concerning fair information practices." (The others are notice, choice, security, and enforcement.) See Fair Information Practice Principles, available at <http://www.ftc.gov/reports/privacy3/fairinfo.htm>; see also FED. TRADE COMM'N, PRIVACY ONLINE: FAIR INFORMATION PRACTICES IN THE ELECTRONIC MARKETPLACE, 106TH CONG. (2000), available at <http://www.ftc.gov/os/2000/05/index.htm#22>.

82. Note that the right of access usually discussed in the privacy context is much narrower in comparison to the right of access that transparency provides, as the former involves only access to information pertaining to the specific individual, rather than everyone else.

83. See Zarsky, *supra* note 3, at 25 n.64 for an explanation of why the Robinson-Patman Act's prohibition of price discrimination does not render this analysis irrelevant.

84. For a discussion of the ability and dangers of discrimination online, see LESSIG, *supra* note 79, at 154; Paul M. Schwartz, *Beyond Lessig's Code for the Internet Privacy: Cyberspace Filters, Privacy-Control, and Fair Information Practices*, 2000 WIS. L. REV. 743, 757.

85. For example, the "lock-in" effect can be achieved by causing consumers to incur high transactional costs when switching vendors, or by creating a comfortable environment (through the use of personal information) that the user will be reluctant to abandon.

86. See Mark Klock, *Unconscionability and Price Discrimination*, 69 TENN. L. REV. 317, 363 (2002) (discussing the various ways such lock-ins can be achieved, both online and offline).

where different segments of the population draw their information from different resources, or shop at different locations. Therefore, we might initially assume that the Internet environment will reduce sellers' ability to carry out such schemes, since other content resources are usually only a click or a *Google* search away. There is no guarantee, however, that the Internet of tomorrow will resemble the one of today. In the future, as the personalization and enclosure of Internet content continues, the web might provide fertile ground for these practices, especially if search engines and other intermediaries will not be able to provide information from the same variety of sources they can access today.<sup>87</sup> In addition, discrimination will pose a serious concern in situations in which all competitors are discriminating in the same way, or in markets with only a few, dominant forces.

In other work, I explain which forms of discrimination are of greater concern than others, divide the possible discriminatory schemes into four categories, and demonstrate the benefits of price discrimination in specific circumstances.<sup>88</sup> The present analysis concentrates on the changes transparency will bring to today's informational status quo, including the present ability to practice discrimination, and whether the creation of a transparent society will mitigate or possibly exacerbate those concerns. I will concentrate the discussion on discrimination in *price*, although many other forms of discrimination may be applicable as well. I start with the good news. Transparency will, to a great extent, protect customers from the detriments of many price discrimination practices. This statement may appear surprising, as at first blush it seems as if a transparent society will actually *enhance* the vendors' ability to discriminate between clients. With transparency in place, the availability of vast amounts of data concerning the locations, interactions, and transactions of individuals will provide additional information about the specific individuals, as well as insight into customers' preferences. Through data mining analysis, vendors can use this information

---

87. For example, the use of mysimon.com or other spider websites will not be as effective, as the vendor websites will learn how to identify the spiders in their search for prices and lock them out. An initial step in this direction is eBay's efforts to block other websites from accessing its datasets. A court recently upheld eBay's efforts, stating that such searches constituted a "trespass" to eBay's "chattel" and therefore could be blocked. *See eBay, Inc. v. Bidder's Edge, Inc.*, 100 F. Supp. 2d 1058 (N.D. Cal. 2000). On this issue, see also Dan Hunter, *Cyberspace as Place and the Tragedy of the Digital Anticommons*, 91 CALIF. L. REV. 439, 483 (2003).

88. See Zarsky, *supra* note 3, at pt. II(1), where I explain at length how those using such price discrimination take advantage of individuals who are lacking in information, are in a hurry, or are just plain lazy. In addition, I discuss situations in which the revenues generated from these practices are used to create a surplus to the firm, rather than redistribute wealth among the consumers (by providing subsidies to others). It should be noted that my current analysis does not specifically address discriminatory practices based on racial or financial parameters. *See id.*, where I explain how these issues should be considered and, to a certain extent, resolved.

to form strong prediction models about consumers' future conduct and thereafter use such insights to the consumers' detriment by discriminating among them on the basis of these results.

However, transparency provides interesting flipsides that will effectively block some of the vendors' abilities to carry out these discriminating schemes. To understand these effects, let us examine the economic prerequisites for a successful price discrimination scheme. As economic analysis indicates, in order for price discrimination to succeed, several conditions must be met: (1) the price set must reflect the buyers' willingness to pay; (2) the seller must obtain market power over the buyer with regard to the specific product; and (3) the seller must prevent or limit arbitrage.<sup>89</sup> Therefore, a competitive market with low transactional and information costs will prevent vendors from overcharging, as competitors will undercut their high prices. Furthermore, the existence of a secondary market, which introduces arbitrage sales of the relevant product between consumers, will discourage such dynamic pricing as well. Transparency accommodates these countering effects through its equal distribution of information, making the practice of price discrimination more difficult.

In a transparent society, vendors can view the personal information their competitors have gathered. With this data, competing vendors can reach out to consumers they suspect are being, or will be, overcharged and locked into higher prices, and offer them competitive deals.<sup>90</sup> They can identify instances where sellers overprice, and then move to interfere in such instances when they can. The competitors' ability to benefit from transparency will depend on the scope of the transactional information made available.<sup>91</sup> Should the information made public include references to the actual products and services purchased and prices paid, competitors will identify the overpriced and "locked-in" consumers with ease. As mentioned above, it is extremely likely that transparency regulation and practices will provide only general information on transactions. However, such data will prove sufficient in enhancing competition and eliminating price discrimination practices, as it can provide enough data to alert competing vendors that specific consumers

---

89. In Zarsky, *supra* note 3, I explain why this dynamic can indeed occur, and in Zarsky, *supra* note 1, pt. IV, I present several other solutions demonstrating how these practices can be inhibited. For the conditions of price discrimination outlined in the text, see Jonathan Weinberg, *Hardware-Based ID, Rights Management, and Trusted Systems*, 52 STAN. L. REV. 1251, 1274 (2000).

90. See Clemens, *supra* note 56, for the overall effects of transparency on a competitive pricing market.

91. Here I refer to *transactional* information as addressed above when describing the hubs of information that the transparent society will provide (Part I.1 above).

might be subject to discrimination by the vendors or marketers they usually use. Thus, transparency allows all parties to lower their information costs<sup>92</sup> and thus promotes efficiency.

Transparency can also mitigate or inhibit discriminatory practices by providing consumers with data about their fellow customers. The existence of information about interactions and transactions in the public domain will set the foundations for an arbitrage market, the existence of which inhibits certain practices of price discrimination. In other words, when customers suspect they are receiving a better price than others, they will seek out those overcharged clients and offer them the same product for the lower price they purchased it for, with an additional marginal markup as their profit. In many instances, the price will still be lower than the one the relevant discriminating vendor is offering. As I explain elsewhere,<sup>93</sup> in today's (and to a greater extent tomorrow's, with an enclosed web looming)<sup>94</sup> information environment, it is quite difficult to formulate such secondary markets, as consumers lack information about who their fellow consumers are, especially when the interactions are in virtual space, where the user interacts solely with the website operator. However, in a transparent society, the availability of information about locations, interactions, and transactions of other consumers will lower transactional costs and facilitate arbitrage sales.

Perhaps most importantly, transparency will add accountability to both online and offline markets and thus deter discriminatory practices to which the public would object. The government and watchdog groups will have an easier task in tracking vendors who are applying such problematic pricing dynamics and in contacting potential victims to warn them of these actions. Therefore, vendors will be reluctant to apply discriminatory schemes from the outset.<sup>95</sup> They will also refrain from practicing discrimination when such actions are prohibited by specific regulation.

*Can there be too much sunlight?* A possible shortcoming of the transparency-based solution in this context is not its inability to confront the fears of discrimination, but in its ability to solve the price discrimination problem *too well*. Even though price discrimination has a deroga-

---

92. Information costs will prove lower for vendors who will have additional information about other consumers, and to the consumers themselves in view of the attempts of other firms to lure overpriced consumers and provide them with additional information.

93. See Zarsky, *supra* note 1, at 54.

94. See *supra* note 87 for a brief explanation as to how such "online enclosure" might arise.

95. This additional accountability factor is important, as it will assist in defeating unfair discrimination in inalienable and personal goods and services, such as insurance. Here, arbitrage and secondary markets will be of no use, but the accountability that transparency accommodates will be of great assistance.

tory sound, it is helpful in many circumstances. Price discrimination schemes should not be understood only as means to generate producer surplus, but in certain situations serve consumer goals by overcharging one group of consumers as a way of subsidizing others who are unable to pay a higher price.<sup>96</sup> From the economist's perspective, in many settings these pricing dynamics are desirable, as they provide every consumer with a price that is in accordance with his or her specific demand curve.<sup>97</sup> Moreover, through these practices, products become accessible to a larger segment of the population, which has been shut out of the market when faced with a high uniform price.<sup>98</sup>

These discrimination dynamics require, at times, a minimal level of opacity. If transparency is to be adopted, full information about the practices of both vendors and consumers will become available, and price discrimination becomes difficult. For instance, those destined for overpricing in a non-transparent world could obtain additional information and thus take advantage of lower prices with ease rather than settle for the higher prices they were previously charged. A shift to transparency will create advantages for sophisticated buyers who may use their newly acquired knowledge to pay the same price as the less privileged consumer, and avoid instances in which they were required to share their wealth by providing a subsidy. Therefore, full transparency will lead to a market in which non-homogenous groups are charged a single price — a detrimental outcome to the weaker segments of society.<sup>99</sup>

A powerful example is airfare pricing in the Internet age. In the pre-Internet era, commercial airline travel provided an excellent opportunity to implement price discrimination. In many cases, three people sitting in the same row all paid different prices for their tickets. Today, however, airline rates are to a certain extent transparent and available on the Internet. Not surprisingly, airlines are already complaining that they are no longer able to reap higher profits from business travelers, who were traditionally charged a higher rate for their short-term/short-notice

---

96. See Klock, *supra* note 86, for common examples of price discrimination, such as student and senior citizens discounts in theaters and public transportation, that in a way are subsidized by the full fare that vendors and service providers charge the rest of the public.

97. See Murphy, *supra* note 42, at 2385. Regarding the benefits of such price discrimination, see Weinberg, *supra* note 89, at 1274.

98. See William W. Fisher III, *Property and Contract on the Internet*, 73 CHI.-KENT L. REV. 1203, 1239 (1998) (making this point with regard to informational goods).

99. Another possible result of transparency with regard to the practices of dynamic pricing is the formation of a great variety of niche markets and personalized products. This outcome will also lead to consumer detriments, as it becomes extremely difficult for consumers to compare products of somewhat similar features and different prices, thus raising transactional costs and again facilitating price discrimination.

business trips. In a tight economy where all businesses are searching for ways to cut costs on travel, business travelers and other sophisticated clients are reading through the airlines pricing strategies and gaining access to the cheaper rates through the web.<sup>100</sup> This causes the airlines to incur losses, and eventually will result in raising airfare, or decreasing service, for all customers. Thus, the common coach traveler will have to bear the costs and pay a higher price to offset the business travelers' new immunity from dynamic pricing.

The conclusion of this analysis of the effects of transparency on discrimination is that society must strive for a proper balance in its regulation of discriminatory schemes, which are made possible in view of the abundance of personal data now available. On the one hand, society should avoid the outcomes that result in abuse, or the creation of an excessive surplus to the vendors; on the other hand, it should still enable forms of price discrimination that promote social equality. Transparency will not provide such a balance, and we must search elsewhere for other solutions that meet this balancing requirement. In a society that is interconnected in so many ways, perhaps too much sunlight — or transparency — is not the best solution.

(D) MANIPULATION AND THE “AUTONOMY TRAP”

The availability of personal information, and the use of data mining tools in its analysis, will confront society with an additional problem — the fear of manipulation and impairment of personal autonomy, which I refer to elsewhere as the “autonomy trap”.<sup>101</sup> To illustrate this problem, let us observe the following fictional example, which introduces a disturbing picture of tomorrow's, and to a certain extent today's, online world and its opportunities:

*In a move befitting his mini-midlife crisis, Mark, 55, has decided to fulfill his childhood dream and buy a motorcycle. Others have anticipated Mark's decision for some time now and are taking action to that effect. Both Portal.com, Mark's usual portal and search engine, and the Gotham Times, the online daily paper Mark reads religiously every day, have predicted Mark's interest in motorcycles with almost pinpoint precision. They have been tracking his actions for some time now — both*

---

100. On these issues, see, e.g., Saul Hansell, *Fare Idea Returns to Haunt Airlines*, N.Y. TIMES, Oct. 27, 2002, at C1. It should be noted that the airlines are in trouble regardless of this factor. See also Klock, *supra* note 86, at 361 (describing the ways that airlines carry out price discrimination).

101. In Zarsky, *supra* note 3, pt. II, I explain this concept in great length. This concept originates from several articles authored by Paul Schwartz (see, e.g., P. Schwartz, *Internet Privacy and the State*, 32 CONN. L. REV. 815, 821 (2000)); Schwartz describes one of the traits of this trap as the “reduced sense of the possible.” *Id.* at 825.

*online and offline — by using their own means of collection as well as data acquired from third parties. The information and knowledge they have gathered (data mining tools reveal that Mark fits the pattern of those who might be interested in a new motorcycle) have moved them to action.*

*Portal.com formed a long-term affiliation with the Darly-Havinson motorcycle company, has a vested interest that Mark will purchase a Darly, and takes action to that effect. As they know from his frequent hotel bookings that Mark travels a great deal, they provide him with information about the great mileage per gallon the Darly offers. They also have some insight into Mark's age and emotional state, and they provide him with advertisements and news articles that capitalize on his insecurities.*

*On the other hand, we have the Gotham Times. It too has picked up on Mark's interest. In view of its financial interests in the Lord car company, the Times will do all that is in its power to convince Mark to abandon his dream and instead purchase a Lord sedan. Therefore, they make sure that his specially tailored newsletter frequently includes articles that outline the dangers of motorcycles, and the higher rate of accidents and deaths involving bikes. They also present him with an ad they believe will capture Mark's attention, as it is of the same font and color as most of the ads that caught his attention in the past, and is presented by his favorite movie star.*

*Mark, who might be vaguely aware of the information constantly collected about him, does not fully grasp the extent of the manipulation attempts constantly occurring in the background. At the end of the day, he makes his decision — but was it really his choice?*

As this still fictitious example demonstrates, the mixture of several novel elements produces a result that might prove harmful to our personal autonomy.<sup>102</sup> These elements are (1) information providers' enhanced ability to garner personal data about their users, especially regarding their interests, preferences, and possible vulnerabilities; (2) their ability to analyze such data in an automatic and efficient manner through the use of data mining applications; and (3) their capability to reach out and provide every user with a personalized package of content,

---

102. For an explanation and analysis of the meaning of autonomy at this juncture, see Zarsky, *supra* note 3, at 40, n.107. In this context, "autonomy" is best defined as a "second order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or to attempt to change these in light of higher-order preferences and values." G. DWORKIN, *THE THEORY AND PRACTICE OF AUTONOMY* 20 (1988). Raz provides a similar definition: "the vision of people controlling to some degree, their own destiny, fashioning it through successive decisions throughout their lives." JOSEPH RAZ, *THE MORALITY OF FREEDOM* 369 (1986).

especially by making use of the Internet's unique infrastructure. Content providers will not only tailor their content to the specific individual upon delivery, but can constantly assess the effectiveness of their marketing schemes and persuasion attempts through the new and updated personal data streaming in from relevant users. Thus, they can create a feedback loop for every user — with the ability to constantly change the information they provide, until they achieve an optimal outcome. These abilities lead to the enhanced opportunity to unfairly persuade and manipulate — a power vested in the hands of a few — and raise concerns both in the context of commercial advertising and agenda-setting by mass media editors.<sup>103</sup>

Transparency, as an overall solution, does not tackle this problem directly.<sup>104</sup> Moreover, as with most of the privacy concerns mentioned, the transparent society provides content providers with additional personal information, which they can use to manipulate the public even further. The ability to carry through these manipulations does not stem solely from the asymmetrical distribution of personal information in society, which is cured by the equality element that transparency fully addresses, or even from lack of accountability. These problematic practices result, among other reasons, from the content providers' control over the information their users are receiving and their ability to tailor such content to every individual. Yet not all is lost, as other traits and effects of the transparent society may offset the dangers of these forms of advanced manipulation.

Transparency might indirectly affect the effectiveness of such manipulations by inducing competition in the media and content provider markets.<sup>105</sup> With information about locations, interactions, and transactions of individuals available in the public domain, new voices will have an easier task in breaking into the tightly held and controlled media market and finding their way to attentive ears. Thus, large media networks will lose some of their advantages as incumbents in the content market, new media entities will have a better chance of surviving in this competitive environment, and the users will receive a larger selection of materials and opinions.<sup>106</sup>

---

103. See Zarsky, *supra* note 3 for a broad description of these elements and the online vendors' newly acquired abilities.

104. Brin neglects to address this issue in his book, and I therefore can only speculate as to how he or others would have thought it should be resolved.

105. In a way, transparency provides a subsidy to new startup content providers by providing them with personal information about potential costumers. I address this issue further in Zarsky, *supra* note 1, pt. II(1).

106. This argument is clearly neglecting an important problem in today's Internet environment — spam and the effects of mailings that are of almost no cost to the sender. If transparency would be put in place today, we may all be flooded by countless solicitations, all striving to meet our

The fact that every individual will be receiving content from a variety of sources can mitigate the detriments arising from manipulative uses of personal information by content providers, as an antidote to powerful means of persuasion is counter-persuasion.<sup>107</sup> In terms of the philosophical and legal literature addressing autonomy, the possibility of receiving competing forms of content protects the users to a certain extent from two possible external constraints on autonomy: the limiting of options, and manipulation.<sup>108</sup> First, generating a variety of voices directly confronts the problematic outcomes of exposure to only limited options (by confronting more voices, the individual is informed of new options). It also provides a partial response to fears of manipulation: When individuals receive content from several sources, they are better equipped to carry out their decision-making in an unbiased manner.<sup>109</sup> Having several sources to draw upon assists individuals in gathering a sufficient amount of information, establishing the accuracy and authenticity of such data, and distinguishing the important factors from those that are of lesser importance.<sup>110</sup>

To demonstrate this point, let us return to the motorcycle example above. With the hypothetical set in a non-transparent society, which includes only limited voices, Mark is subject to short but powerful interventions. In this example, those few entities with access to both Mark's eyeball and his personal information are clearly at an advantage and stand a better chance of influencing him. However, in a transparent society, where many content providers have access to Mark's personal

---

needs. These solicitations will also come from vendors and marketers that are offering products that only remotely match our requirements — but the low costs of mailing make such solicitations profitable. However, I believe this problem is manageable, and various solutions will be put in place to convince marketers to send out emails or other messages only when it is economically viable and the chances that we are indeed interested are high. For example, this can be achieved by charging a nominal fee for every email after sending a certain number of emails (thus convincing mass emailers to reconsider their actions). For examples of such solutions, see Brad Templeton, *Proper Principles for Challenge/Response Anti-Spam Systems*, at <http://www.templetons.com/brad/spam/challengeresponse.html>; Microsoft Research, *Penny Black Project*, at <http://research.microsoft.com/research/sv/PennyBlack>.

107. On this issue, see David Strauss, *Persuasion, Autonomy, and Freedom of Expression*, 91 COLUM. L. REV. 334, 347 (1991). Strauss discusses the various forms of persuasion and how they might be effectively countered. Note, however, that at a later stage of the argument, Strauss concedes that the effects of counterpersuasion are somewhat limited.

108. See Richard H. Fallon, Jr., *Two Senses of Autonomy*, 46 STAN. L. REV. 875, 886 (1994), for an analysis of various forms of constraints on autonomy.

109. Strauss, *supra* note 107, at 362, discusses the ways a speaker can manipulate listeners, mentioning the (a) withholding of opposing arguments, (b) distorting the truth in ways that are not outright false, or (c) appealing to the listener in a subliminal way. These actions can be understood as referring to different parts of the decision-making process I mention in the text.

110. For an in-depth description of the decision-making process and how the autonomy trap interferes and erodes it, see TAL ZARSKY, *DISARMING THE TRAP, DEFINING, EXPLAINING, AND SOLVING THE CHALLENGES OF THE AUTONOMY TRAP* (Working Paper, on file with author).

information and possibly to Mark as well, he will be subject to suggestions and manipulations from a wide array of entities, each attempting to tilt Mark's opinion and behavior in their direction. For example, in this new environment, one party's attempt to hide specific forms of information will be offset by another party's reference to the hidden information. These effects will create an eventual balance in the information Mark receives, and allow him to reach a more autonomous decision.<sup>111</sup>

The picture painted above might be too optimistic. Equal access to personal information may be insufficient in promoting competing voices in the content market. Therefore, even a transparent society may require additional regulation and enforcement to ensure diversity.<sup>112</sup> Such regulation must also ensure that incumbents do not abuse their market power and that other content providers will have an opportunity to carry their content to users.<sup>113</sup> Furthermore, the availability of balanced content does not always shield individuals from all forms of manipulation.<sup>114</sup> External entities can still attempt to manipulate and thus impinge on the individuals' autonomy by interfering with their decision-making process.<sup>115</sup> This can be done, for instance, by creating "noise" to interfere with a rational cognitive process or by invoking the individual's subconscious or emotions, by injecting irrational factors into an otherwise rational equation. In our motorcycle example, the content providers use manipulating interventions and persuasive advertisements to induce Mark to make a decision premised on emotion or other irrational factors — for instance, by playing on his insecurities and using advertisements that previous analysis proved to be highly effective.<sup>116</sup>

---

111. Autonomy, especially in this context, is a matter of degree. See Fallon, *supra* note 108, at 877. In this context, I therefore submit that transparency will facilitate a suitable level of autonomy for users and consumers when interacting with content providers.

112. Note that the FCC has recently decided to relax the limits on media concentration. See Stephen Labaton, *Deregulating the Media: The Overview; Regulators Ease Rules Governing Media Ownership*, N. Y. TIMES, June 3, 2003, at A1. Note that these rules created a public uproar and at this time are blocked by Congress and the courts. See Ron Orol, DCC Blocked on Merger Rules, THE DEAL, Mar. 9, 2004, at <http://www.thedeal.com/NASApp/cs/ContentServer?page name=IWM&c=TDDArticle&cid=1078420952031>.

113. In addition to the problems addressed in the text, it is also possible that every one of the various forms of media will be providing the same form of content and that certain forms of important content will be silenced (for example, in the motorcycle example, the voice of the Green Party, which would advocate walking or taking the bus). It is very difficult to predict the outcome of transparency-based regulation on this market or to predict whether the creation of many voices will solve the problems addressed above. On this issue, see Yochai Benkler, *Siren Songs and Amish Children: Autonomy, Information, and Law*, 76 N.Y.U. L. REV. 23, 68 (2001). Benkler mentions several reasons for the failure of this "market-will-solve-it" hunch, including high transactional costs and negotiation costs for the individuals.

114. Regarding the elements of the decision-making process, see note 110 and accompanying text.

115. *Id.*

116. See Zarsky, *supra* note 3, pt. II, for a demonstration and explanation as to how data

The simplest and most apparent response to this set of autonomy concerns is that they will be resolved through social adaptation. It is indeed possible that individuals, and society in general, will adapt to the manipulative reality of a transparent society, or one that facilitates these advanced means of persuasion. As they have in the past, our minds and senses will learn to manage information overloads and attempts to persuade, and personal autonomy will be restored. However, this forecast might be too optimistic, as the quantity and quality of the tailored eye candy and personalized manipulation that we will confront in a transparent society will be unprecedented, in view of the personal data now available.<sup>117</sup> Since transparency does not confront these issues, additional regulation will be needed in a transparent world. For instance, new rules must require content providers to notify those subject to manipulations that personal information is being used to tailor the content they are receiving.<sup>118</sup> However, in a transparent society, such notification might still prove insufficient in overcoming manipulative measures. Therefore, before considering a transparency-based solution, research is needed to establish the effects of manipulations in a situation that resembles a transparent society; i.e., where a vast amount of information about individuals' locations, interactions, and transactions is available. Such research must take into account the existence of data mining applications that can now provide additional insights on the basis of sketchy data, as well as the Internet's ability to facilitate both the collection and delivery of specific content to each user. Should research indicate that transparency indeed facilitates manipulations that are too powerful to be countered by notification, this solution must be rejected. Preserving autonomous thought clearly outweighs the benefits transparency provides.

---

mining tools and analysis can be used to find the most effective means of manipulation for every consumer and user.

117. Even in today's world, people are finding it difficult to overcome the information loads they are facing when making decision. For example, experiments have shown that too many options confuse customers and make it very difficult for them to reach decisions. See Stephen Dubner, *Calculating the Irrational in Mathematics*, N.Y. TIMES, June 28, 2003, at B7, for examples of relevant experiments, including experiments conducted in grocery stores, where customers were provided with varying numbers of products.

118. For the importance of "knowledge of manipulation," see Strauss, *supra* note 107, at 363 (describing the balancing effects of knowledge in the event of manipulations). Elsewhere I address the extent and importance of such notification (*see* note 110 above). Such notification requirements should consist of a duty to inform users that they are receiving content that was (1) selected (2) especially for them, as part of a group or as individuals, (3) on the basis of specific personal information. With such knowledge in hand, the user obtains the ability to manage and counter these manipulations, thus reaching an autonomous decision.

### I.3. *Transparency as “Leveling the Playing Field”?*

Providing solutions to the fears and problems of today’s information environment does not justify a shift to transparency, since the objectives can be achieved by less controversial means.<sup>119</sup> A shift to a transparent society also surely strives to promote equality in the information market. Can transparency actually achieve this optimistic objective? I examine this assumption by concentrating on the information balance between sellers and buyers in the marketplace.

At first glance, one can assume that a transparent society will be successful in promoting equality. This new society will provide all citizens with similar information resources. Therefore, vendors and content providers will not be the only ones with the ability to use personal and business information in sorting potential customers, constructing pricing schemes, and possibly preying on the even temporary vulnerability of consumers. In a transparent society, consumers will also have access to the collected information and could use such data in various ways. For example, buyers can review information about vendors and corporations and use it in search of sellers who are in dire need to sell. Moreover, in a transparent society, customers can engage in “offensive data mining” by using data mining to analyze the actions of commercial entities and predict their future dealings. Consumers may carry out these analyses themselves, or with the assistance of intermediaries that will process their requests. Currently, most data mining intermediaries — such as DoubleClick, which uses data mining analysis to match advertisements to users’ personal profiles — accommodate vendors and marketers that are, by no coincidence, today’s dominant collectors. Transparency and the availability of personal data on a grand scale hopefully will promote the emergence of other intermediaries that will act to the benefit of private customers as well. Such agents will compare prices at various outlets<sup>120</sup> and seek out patterns of consumer sales, or assess the vendors’ revenues and try to establish how much leverage the sellers have in each transaction.<sup>121</sup> The automatic nature of data mining will allow these analyses to be carried out on a large scale and with relative ease. More-

---

119. For examples and an analysis of appropriate solutions, see Zarsky, *supra* note 1, pt. IV.

120. Information of this sort is now available via today’s online intermediaries such as <http://www.mysimon.com>.

121. However, there might be some difficulties with regard to offensive data mining used to predict the behavior of firms. In general, corporation-based predictions could be achieved by collecting data pertaining to specific corporations, analyzing it, and then making educated guesses as to future actions and outcomes. However, unlike individuals, corporations might not demonstrate consistent patterns of behavior over a long period of time. The management and board of a corporation may change, and the future actions of their successors will be unknown and unpredictable.

over, the Internet's infrastructure will also facilitate equality by providing accessibility and enabling simple searches in the vast amounts of data to be provided online. Given these abilities, the prices and services that sellers quote and provide to specific customers will reflect the additional knowledge the buyers can obtain. Thus, the market will adjust, and the effects of asymmetric information in today's information market will disappear.<sup>122</sup> At a later stage, this might even lead to the evolution of a different form of marketplace. Future online, or even offline, retail transactions may consist of automatic negotiations between intermediaries representing buyers and sellers in what resembles today's online auction market. Each party will provide his or her virtual negotiator with personal information, preferences, and guidelines for negotiations, and allow the automatic negotiation to proceed. Such practices offer significant advantages over today's dynamics and allow consumers to effectively signal to sellers their general interest in a transaction as well as their reservations about the quoted price.<sup>123</sup>

However, as I demonstrate below, this ideal reality of equal access and data analysis will remain utopian, and transparency will fail in promoting equality.

#### LEVELING VS. MINING

An additional look at the outcomes of a transparent society leads to the conclusion that it will prove insufficient in leveling the playing field between the stronger and weaker players in the current information market. The ambitious objective of achieving equality by promoting transparency faces fundamental difficulties in light of the public's capability to use the data that transparency provides and the ability to effectively conduct data mining analysis. In today's technological reality, equal access to information is insufficient, and access to raw data is almost as good as having no access at all. To grasp and analyze the vast amounts of information available, sophistication is now the key. Sophistication will remain unequal in a transparent society,<sup>124</sup> as a large segment of the population will face both the lack of tools and the lack of ability, time, and knowledge. I address the importance of such sophistication by dem-

---

122. With regard to the asymmetries in today's information market, see Zarsky, *supra* note 1, at 40-41; see also Michael Froomkin, *The Death of Privacy?*, 52 STAN. L. REV. 1461, 1502 (2000).

123. This description somewhat resembles the automatic negotiations made possible in the privacy market by agents such as P3P applications — a solution advocated by LESSIG, *supra* note 79, at 160; see also Zarsky, *supra* note 1, pt. I.

124. See Daniel J. Solove, *Privacy and Power, Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393, 1455-56 (2001).

onstrating how inequality will persist in a transparent society in view of unequal access to and use of data mining applications.

The practices of “offensive data mining” that I address above can be pursued only if individuals have access not only to information but also to data analysis technology. However, data mining applications are expensive and at times beyond the reach of the general public. In addition, the newest and most sophisticated tools might be patent protected by their developers, who might strategically choose to place them only within the reach of large companies, rather than individuals and small consumers. Should such uneven access persist in a transparent society, transparency will in fact increase the disparity between individuals and large entities, rather than level the information playing field.<sup>125</sup> Since a large segment of society will be denied access to the real knowledge that transparency can provide, the promise of equality via transparency is an empty one. A shift to a transparent society might lead to the emergence of a new generation of data mining applications that are cheap, accessible to the public, and operated by intermediaries with the interest of serving the private customer. However, one can only speculate as to the chances of this optimistic outcome, which perhaps could be encouraged by regulation.

In addition, access to data mining applications takes us only half-way toward equality. Using the data mining applications and analyzing their results is a difficult task, which will not be equally mastered among the various segments of society. Even with equal access to both the information and the analysis tools, the disparity between individuals and firms as well as among the individuals themselves will persist, due to the disparity in their ability to use the results of data mining applications. The data mining tools, even the simplest ones, are extremely complicated to apply, and their benefits heavily depend on the knowledge and experience of the user. As a result, those less privileged will again find themselves in an inferior position. To counter such unequal outcomes, transparency must be part of a broader policy scheme that will include research initiatives and funding toward the development of simpler data mining applications, as well as scholastic efforts to enable a broader segment of society to benefit from the information made available.

A final threat to equality in the transparent society would result from the actions of those purposefully excluding themselves from transparency. Individuals with sufficient resources will take appropriate measures to conceal their actions and movements, thus circumventing the transparency regime. These members of society will become free

---

125. See Kang, *supra* note 10 at 124, who makes a similar argument with regard to the distribution of anonymizing applications. (I discuss this issue at length below.)

riders, benefiting from the vast opportunities of open information without exposing themselves to the risks of transparency. It is most likely that vendors and commercial entities will be those with the ability to conceal their actions, thus leading us back to a world that resembles today's one-way mirror. A shift to a transparent society will require regulation addressing this issue directly, but any scheme will still allow for exceptions and free riders.<sup>126</sup>

Transparency, therefore, which may have seemed an egalitarian and even socialistic solution, will most probably lead us back to the familiar imbalance of power between collectors and collectees: Those equipped with the better resources as well as the benefits of education and experience will make the most of the available data, while others are faced, or even smothered, with vast amounts of unusable information.

Concluding my broad analysis of the transparent society, I establish that the transparency concept requires and is worthy of additional research and exploration. More specifically, I conclude that revealing information about the locations, interactions, and transactions of the general public will not necessarily level the playing field between the stronger and weaker players in the information society. In a transparent society, inequality persists as data mining and other analysis tools create a disparity between real knowledge and raw data. Furthermore, it is unclear whether the transparency solution can actually resolve the problems stemming from the use of personal information, especially with regard to its impact on personal autonomy and the possible abuses of personal information.

## II. ANONYMITY

### II.1. *What Is Anonymity?*

Thinking outside the box about solutions to today's privacy concerns leads us directly to the anonymous society. In many respects, anonymity is the opposite of transparency, where instead of increasing the flow of meaningful personal information, it brings it to an immediate halt. Anonymity solves today's problems by permanently disconnecting the information collected from the individuals to whom it pertains. Recent technological innovations can facilitate a switch to anonymity, both online and offline. But anonymity comes at a high price and creates problems that can only lead to the failure of this proposed solution.

Unlike transparency, the meaning and structure of an anonymous

---

126. One model suggested by Brin is based on seclusion; namely, those that will not provide information are banned from participating in the information sharing. *See supra* note 17 above and accompanying text. It is not clear, however, how this could be carried out if all information will be open.

society is quite clear and is frequently addressed in the legal literature, both in case law<sup>127</sup> and secondary sources.<sup>128</sup> Anonymity is usually mentioned in the context of the right to read, write, speak, and distribute content without exposing the identity of the relevant individual. In this context, the argument for anonymity has a constitutional aspect as well.<sup>129</sup> The Supreme Court has found on several occasions that the First Amendment protects a right to anonymous speech.<sup>130</sup> Other cases in lower courts have specifically addressed the right of anonymous speech online, noting the importance of such speech in the Internet environment.<sup>131</sup> The right of anonymous speech was protected in such situations: as circulating leaflets,<sup>132</sup> soliciting,<sup>133</sup> and promoting a petition.<sup>134</sup> However, a recent ruling indicates that perhaps the constitutional protection of anonymous speech should be confined to instances where its uses safeguard those who fear retaliation or an unwanted intrusion of privacy.<sup>135</sup>

---

127. See *infra* notes 130-135.

128. For a discussion and analysis of anonymity in the Internet society, see generally A. Michael Froomkin, *Flood Control on the Information Ocean: Living with Anonymity, Digital Cash, and Distributed Databases*, 15 J.L. & Com. 395 (1996). Note, however, that this article emphasizes the importance of anonymity with regard to free speech and censorship. For additional sources, see LESSIG, *supra* note 79, at 139-40; Julie E. Cohen, *A Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 CONN. L. REV. 981, 1003-38 (1996).

129. For the benefits of anonymity in general, see Ann Wells Branscomb, *Anonymity, Autonomy and Accountability: Challenges to the First Amendment in Cyberspaces*, 104 YALE L.J. 1639, 1642 (1995).

130. *Watchtower Bible & Tract Soc'y of N.Y., Inc. v. Vill. of Stratton*, 536 U.S. 150 (2002). Additional cases discussing the right of anonymous speech are *Buckley v. Am. Constitutional Law Found.*, 525 U.S. 182 (1999) and *McIntyre v. Ohio Elections Comm'n*, 514 U.S. 334 (1995).

131. *Doe v. 2TheMart.com, Inc.*, 140 F. Supp. 2d 1088, 1097 (W.D. Wash. 2001); *ACLU v. Johnson*, 4 F. Supp. 2d 1029, 1033 (D.N.M. 1998), *aff'd*, 194 F.3d 1149 (10th Cir. 1999); *ACLU of Ga. v. Miller*, 977 F. Supp. 1228, 1230 (N.D. Ga. 1997). On issues of encryption, see *Junger v. Daley*, 209 F.3d 481 (6th Cir. 2000); *Bernstein v. U.S. Dep't of Justice*, 176 F.3d 1132 (9th Cir. 1999) (withdrawn pending rehearing *en banc*).

132. *McIntyre*, 514 U.S. at 334.

133. *Watchtower Bible*, 536 U.S. at 150.

134. *Buckley*, 525 U.S. at 182.

135. See *In re Verizon Internet Services, Inc.* 257 F. Supp. 2d 244, 259 (D.D.C. 2003) (also citing *Watchtower Bible*, 536 U.S. at 166), where the court discusses the anonymity rights (or lack thereof) of individuals swapping songs online. Note, however, that this case has been recently remanded and vacated by the Circuit Court in *RIAA v. Verizon Internet Services Inc.*, 351 F.3d 1229 (D.C. Cir. 2003). However, in this later case, the Circuit Court did not address the anonymity rights of online users, but addressed the District Court's understanding of specific provisions of the Digital Millennium Copyright Act (DMCA) of 1998, Pub. L. No. 105-304, 17 U.S.C. §§ 1201-05. Therefore, the District Court's statements about these issues are worth mentioning, and might be echoed in future rulings. For a similar analysis of the two Verizon cases, see Sonya K. Katyal, *The New Surveillance*, CASE W. RES. L. REV. (forthcoming 2004), available at [http://islandia.law.yale.edu/isp/digital%20cops/sonia\\_katyal.pdf](http://islandia.law.yale.edu/isp/digital%20cops/sonia_katyal.pdf). (regarding the trial court's understanding of anonymity) and *id.* at 3 (with regard to the D.C. Circuit's ruling).

For this Essay, I suggest we paint a picture of an anonymous society using a much broader brush. Anonymity in the context of this discussion should not only shield individuals while speaking or reading, but also when roaming about, interacting, and transacting. Therefore, I suggest we consider anonymity-based solutions that provide individuals with disposable, one-time identities that cannot be traced back to their actual selves, or that allow them to avoid the eye of any data collector. The simplest means of facilitating anonymity is the use of cash, as opposed to credit or other smart card devices. Payment with cash does not require verification of credit or identity, and the name of the purchaser is not registered. Thus, no link is formed between the participant in the transaction and his actual identity or other patterns of behavior. Cash, therefore, is the basic “anonymizer.”

Anonymity-creating tools may seem vague and detached in real space, but they are easily fleshed out in cyberspace, where surveillance is carried out on a much higher level. Individuals can conceal their online locations by using software tools that assign random and temporary identification numbers for every Internet session, thus leaving no link to the actual user who visits websites or chat rooms. (Similar applications are now appearing in the brick-and-mortar world as well.)<sup>136</sup> Individuals can conceal their interactions through the use of re-mailing applications<sup>137</sup> that hide the identity of senders and the source of email messages.<sup>138</sup> They also can protect the secrecy of their communications by using encrypting tools to secure the content of messages. Finally, users of e-cash<sup>139</sup> can carry out e-commerce transactions without leaving a clue as to their true identities.

Taking these applications into account, it is clear why anonymity as an overall solution is “outside the box” when compared to the solutions that fit within the “information flow paradigm” I addressed above; namely, solutions that regulate privacy concerns by addressing the collection-analysis-implementation stages of the personal information flow. Nonetheless, anonymity-based solutions do not restrict the collection of

---

136. See John Markoff, *Protesting the Big Brother Lens, Little Brother Turns an Eye Blind*, N.Y. TIMES, Oct. 7, 2002, at C3 (discussing the protection from surveillance cameras by pointing lasers at them).

137. For a description of such technology, see Kang, *supra* note 10, at 1243. See also Branscomb, *supra* note 129, at 1661.

138. On the issue of anonymous interactions, see Schwartz, *supra* note 10, at 1628, where he discusses the case of U.S. Navy Senior Chief Timothy R. McVeigh, in which AOL provided the Navy with information linking McVeigh to the email alias “boysrch.” The Navy sought to discharge McVeigh because of his sexual orientation. “Boysrch” had identified his marital status as “gay.”

139. For a model of e-cash, see David Chaum, *Achieving Electronic Privacy*, SCIENTIFIC AM., Aug. 1992, at 96-101.

personal information, which could carry on as before but would lead only to the accumulation of meaningless data. The analysis of the information is not regulated directly, but becomes futile nonetheless, as the recorded actions and transactions cannot be linked to specific individuals or prior occurrences. Thus, in an anonymous world, marketers and information collectors will be forced to limit their data analysis to information their customers voluntarily provide<sup>140</sup> and to sales-related information they record during their ongoing operation.<sup>141</sup> Finally, it will be impossible to use personal data to the same extent it is used today, not because of limitations on such actions, but because of the lack of relevant information.

Should society broadly adopt anonymizing tools and form an anonymous society, the privacy problems I address above will quickly melt away. As an anonymous society will not facilitate the formation of the three hubs of personal information (location, interaction, and transaction), only a limited amount of data could to be abused. The collectors' databases will be filled with meaningless data, thus mitigating fears of errors. Similarly, with no personal information about users and customers, the elaborate practices of discrimination and manipulation described above are rendered impossible.

## II.2. *The Troubles of the Anonymous Society*

The anonymous society will lead to the disappearance of many of the privacy problems addressed above but in turn will generate new concerns. This solution will also face powerful adversaries, who will undermine any overall shift to anonymity.

First, there are the practical problems. A successful shift to anonymity depends on the public's acceptance of the anonymizing technology. However, the anonymity-enhancing tools are still cumbersome, and are used almost exclusively by the computer savvy,<sup>142</sup> while the

---

140. At this point, a possible powerful critique of the anonymity solution arises: Since collectors have an overwhelming advantage in understanding the benefits of personal information, they will be able to convince the consumers to submit the information voluntarily, thus circumventing anonymity. To this powerful claim, I respond that, with a shift to anonymity, the entire perspective as to the norms of information collection will shift as well, making it unreasonable for individuals to give up on their anonymity. However, I concede that this point is perhaps the Achilles's heel of the entire analysis.

141. For example, information about what products were sold and at what times, but without the ability to establish which customer purchased what product.

142. See, e.g., Charles C. Mann, *A Primer on Public-Key Encryption*, ATLANTIC MONTHLY, Sept. 2002, at [http://www.theatlantic.com/issues/2002/09/mann\\_g.htm](http://www.theatlantic.com/issues/2002/09/mann_g.htm). This article quotes an experiment carried out at Carnegie Mellon University in 1999, when 12 experienced computer users were requested to send and receive PGP encrypted messages. These users utterly failed in this assignment.

general public is unfamiliar with this technology.<sup>143</sup> Therefore, for a real global solution to succeed, a new generation of anonymizing applications must emerge to enable wide access and use.<sup>144</sup> In addition, these applications would have to be incorporated into the gatekeeping applications widely used — such as operating systems or browsers.<sup>145</sup> Moreover, it will probably take a governmental initiative to lead a substantial portion of society into anonymity.<sup>146</sup> As I explain below, the government will be reluctant to take such a leading role to promote anonymity.

Beyond the practical difficulties, anonymity will come at a high price to society. Anonymity adversely affects society by causing the loss of accountability. Communications in an anonymous environment seem to strip users of the civility that face-to-face encounters engender. Anonymity also facilitates the distribution of false information.<sup>147</sup> In an anonymous society where locations, interactions, and transactions are hidden behind a veil of anonymity, the lack of accountability will spread to all areas of conduct, interfering with the formation of business and other relationships and allowing individuals to act without inhibitions.

In addition to the loss of accountability, shifting toward anonymity leads to a loss of knowledge. The shift to an anonymous society will render the analysis of personal information impossible, thus causing the loss of its beneficial outcomes. The databases will be emptied of any individual-related information, and will be left with only the ability to carry out basket analyses, which provide answers to questions such as which products or services were purchased together, and when. Commercial entities that might benefit from the collection and analysis of personal information will strongly oppose such a shift, and have a valid point in doing so. As I explain elsewhere, the analysis of personal data boosts innovation, assists startup ventures, and allows us to create value

---

143. For example, Paul Schwartz writes of the “blinking twelve” problem — the fact that users are so uncomfortable with new technology that they cannot even set the timer on their VCRs. *See* Schwartz, *supra* note 84, at 752.

144. On this point, Kang, *supra* note 10, points out that the technology enabling anonymity is unfairly distributed.

145. Currently, it is much easier to incorporate an anonymizer in the Linux operating system. *See*, Thomas C. Greene, *Internet Anonymity for Linux Newbies*, THE REGISTER, Aug. 28, 2002, at [http://www.theregister.co.uk/2002/08/28/internet\\_anonymity\\_for\\_linux\\_newbies/](http://www.theregister.co.uk/2002/08/28/internet_anonymity_for_linux_newbies/). As for privacy and anonymity-enhancing issues relating to Microsoft Windows and Explorer, *see* Thomas C. Greene, *Internet Anonymity for Windows Power Users*, THE REGISTER, Dec. 5, 2001, at [http://www.theregister.co.uk/2001/12/05/internet\\_anonymity\\_for\\_windows\\_power/](http://www.theregister.co.uk/2001/12/05/internet_anonymity_for_windows_power/).

146. The overall problem with market-based solutions to privacy concerns is that it is extremely difficult to construct a profitable business model to provide privacy. In general, individuals undervalue their privacy and sell it very cheap. For example, Forrester Research has concluded that the value individuals attach to their total privacy is about \$5 a month. *See* Bob Tedeschi, *Technology Briefing Internet: The Price of Online Privacy*, NY TIMES, June 19, 2002, at C4.

147. On the issue of anonymity and accountability, *see* Branscomb, *supra* note 129, at 1645.

in several ways. Most of these benefits will be lost in a shift to anonymity.<sup>148</sup>

#### ANONYMITY'S ENEMIES

The anonymous society is sure to face a coalition of powerful antagonists besides the information collectors. First, there are the major copyright owners and licensors; namely, the film and music industries.<sup>149</sup> These powerful entities would strongly oppose the enhancement of anonymity, since they fear such tools will render the enforcement of intellectual property rights impossible. With anonymity in place, copyright materials will be passed across the Internet under the noses of the helpless copyright owners, who will be unable to trace and deter such actions.<sup>150</sup> For enforcement of copyright and other forms of intellectual property to take place, information about interactions must be evident, collectable, and lead to a specific physical person.

However, the strongest opposition to an anonymous society will come from the state, which will cite military and security concerns. Applications that enhance anonymity are effective tools in the hands of organized crime members, drug dealers, and terrorists.<sup>151</sup> Law enforcement cannot monitor an encrypted and anonymous net, creating a safe haven for criminals.<sup>152</sup> Wrongdoers can use untraceable communications and anonymous interactions to achieve goals that are clearly detrimental to society.<sup>153</sup> In the past, the U.S. government tried to preserve its ability to track and access encrypted communications<sup>154</sup> as well as communications in general<sup>155</sup> and objected to the export and publication

---

148. For a discussion of the benefits of personal information analysis and the possible loss of such benefits in view of privacy regulation, see Zarsky, *supra* note 1, pt. II(1).

149. The recording industry's concern with Internet anonymity was evident in its recent demand that Verizon hand over the identity of an individual who allegedly downloaded 600 songs using peer-to-peer software. See Amy Harmon, *Verizon Ordered to Give Identity of Net Subscriber*, N.Y. TIMES, Jan. 22, 2003 at C1. See also *RIAA v. Verizon Internet Servs. Inc.*, 351 F.3d 1229 (D.C. Cir. 2003). For an in-depth analysis of such cases, see Katyal, *supra* note 135.

150. Froomkin, *supra* note 128, at 404. Note the distinction between anonymity and secrecy. An anonymous net prevents the tracing and deterrence of infringers. A secret net (promoted by encryption) prevents the revelation of the actual infringement.

151. See AMITAI ETZIONI, *THE LIMITS OF PRIVACY* 78 (1999), for a list of the specific threats that encryption and anonymity create in this context.

152. Froomkin, *supra*, note 128, at 402.

153. For the way terrorist groups made recent use of encryption tools, especially in the context of the 9/11 attacks, see Farhad Manjoo, *Is Big Brother Our Only Hope Against Bin Laden?*, (Dec. 3, 2002), at <http://www.salon.com>.

154. One form of action was to persuade the public to accept a protocol (the "Clipper Chip") that provides the government with a copy of the encryption keys as well. On these issues, see, e.g., Froomkin, *supra* note 122. With regard to governmental back doors, see *id.* at 1488.

155. To maintain its ability to accomplish electronic surveillance, the government has enacted the Communications Assistance for Law Enforcement Act (CALEA) of 1994, Pub. L. No. 103-

of encrypting programs.<sup>156</sup> Although the government has recently eased its regulation of these matters,<sup>157</sup> it will surely stop far short of advocating the implementation of anonymity-enhancing schemes.

In view of the difficulties articulated above, it appears that anonymity will fail as an effective global solution or that it will meet powerful opposition it will not be able to overcome. More elaborate forms of anonymity can prove practical and useful. I now turn to one of them — pseudonymity — as a possible option.

### III. PSEUDONYMITY

#### III.1. *What Is Pseudonymity?*

The third and last “outside the box” solution I address in this Essay is pseudonymity. I refer to pseudonymity as *the use of a “virtual” personality or personalities by one physical individual when interacting in cyberspace or elsewhere*. Thus, rather than using a different face for every transaction, as with anonymity, the pseudonymous user will have several consistent faces that can be used interchangeably and are detached from the user’s physical persona. A shift to a pseudonymous society could mitigate several of the problems caused by the collection and analysis of personal data, but again with unwanted side effects. Although pseudonymity is widely discussed in the Internet context, it is a concept that is far from novel, as writers and political figures have chosen to hide behind the mask of pseudonymity in their publications in previous centuries as well. For example, the fact that several authors of *The Federalist Papers* made use of pseudonyms is often mentioned to emphasize the importance of pseudonymity. Legal scholars discuss the use of pseudonyms in the Internet society in various contexts, only one of which is the enhancement of privacy.<sup>158</sup> For instance, the importance of pseudonymity is often mentioned in connection to the right of free speech and the battle against censorship.<sup>159</sup> In this Essay, I approach pseudonymity from a somewhat different angle — one that is mainly concerned with actions and transactions, rather than forms of speech.

---

414, 108 Stat. 4279, which requires telecommunication providers to help the government in executing various forms of surveillance. See Solove & Rotenberg, *supra* note 47, at 337.

156. See, e.g., *Bernstein v. United States Dep’t of Justice*, 176 F.3d 1132 (9th Cir. 1999) (withdrawn pending rehearing *en banc*). See also Smith, *supra* note 40, at 30.

157. See *id.*; ETZIONI, *supra* note 151, at 101. Also, on these issues, see articles referred to at <http://moglen.law.columbia.edu/LIS/index.html>, Eben Moglen’s “Internet Society” course web page, Columbia Law School.

158. A good resource on these issues is David Post, *Thoughts on Anonymity, Pseudonymity, and Limited Liability in Cyberspace*, 1996 U. CHI. L. FORUM 139.

159. See sources addressed in notes 130 and 131 with regard to the First Amendment right to speak anonymously. These cases apply, to a certain extent, to pseudonymity as well.

## TRACEABLE VS. UNTRACEABLE PSEUDONYMITY

Recent scholarship on the possible forms of a pseudonymous society online distinguishes between two options: traceable and untraceable pseudonymity.<sup>160</sup> Untraceable pseudonymity provides an individual with the exclusive use of an alias, or several aliases, which cannot be traced back to his or her physical persona by anyone or in any way.<sup>161</sup> Such a mechanism allows the speaker to create ongoing accountability and goodwill for the specific alias (which takes on a personality of its own), without revealing its real source and the individual's own identity. However, communications through these means usually do not allow for others to contact the alias directly.<sup>162</sup> The greatest technical problem such schemes entail is the difficulty in ensuring that the designated individual has exclusive use of the alias—in other words, to ensure that imposters do not take over the user's identity. In the past, such exclusivity was maintained through the writer's unique signature or style. Today's technology offers other sophisticated means to meet this objective, such as electronic signatures that ensure that the messages the alias is posting always originate from the same, unknown source.<sup>163</sup>

The fact that the pseudonym and the real person cannot be linked is not always an advantage. When using untraceable pseudonyms, the pseudonym can send out messages, but it will be quite difficult for others to contact him or her directly and privately.<sup>164</sup> This problem makes untraceable pseudonymity an unattractive option as it creates difficulties in transactions that require two-way communications. Thus, I shift my focus to the second form—traceable pseudonymity.

Traceable pseudonymity enables a two-way link between the pseudonym and the physical self by allowing the individual to directly and discreetly receive messages intended for the alias.<sup>165</sup> With two-way communications, a pseudonymous society can accommodate actual

---

160. This framework is suggested by Froomkin, *supra* note 128.

161. *Id.*

162. Direct communications with the alias are not possible, as in this scheme the alias does not provide information as to how it could be contacted directly. Providing such information would allow the speaker's identity to be traced.

163. Digital signatures can in general ensure that the document has been signed by the same individual every time and cannot be forged through the use of cryptographic tools. For a simple explanation of the mechanics of digital signatures, see <http://www.youdzone.com/signature.html>.

164. Others can reach out to the alias in public by posting messages only the alias could read, but such communications are not ideal. However this problem could be addressed through the use of asymmetric encryption: First, the sender can provide the recipient with a public key. Thereafter, the recipient can leave a message for the sender in a public place (for example, a chat room) that is encrypted with the sender's public key. Only the sender can decrypt and read this message. However, this form of communications is complicated and risky—as adversaries could still track those consistently examining the “billboard,” or public space.

165. Froomkin, *supra* note 128.

interactions or business transactions and is a better simulator of the physical environment. In a traceable pseudonymous society, only a few, such as mailers and specific intermediaries, can connect the virtual selves and the physical selves and may do so only under specific circumstances and certain terms. Most of society will not be privy to such identifying information while interacting with other virtual personas and will only interact with the pseudonym.

A traceable pseudonymous society introduces substantial advantages in comparison to the anonymous society. While anonymity requires that we interact with blank faces and constantly meet new people, traceable pseudonymity allows us to interact with consistent personalities, with whom we can transact directly and carry out two-way conversations. Moreover, traceable pseudonyms present opportunities for marketers and content providers as well. Individuals visiting the virtual mall or interacting with content providers can still be subjected to limited profiling that will facilitate a linkage between their current conduct and past patterns of behavior. The collectors and profilers thus maintain the ability to push information, endorsements, and advertisements on the basis of the personal information they collect, the analysis they conduct and the profiles they construct. But — and this is a big “but” — they are unable to create a full profile of all the individual’s virtual pseudonyms, and cannot connect such profiles to his or her real, physical identity.

Traceability, however, is a two-edged sword. On the one hand it creates many benefits; on the other it increases the risk that the physical identity will be revealed and the entire objective of pseudonymity shattered. Not surprisingly, legal controversy arose about the fiduciary duties of the intermediaries who connect the pseudonym and the real identity, and concerning the situations in which the intermediaries are to withhold or disclose information about this critical connection.<sup>166</sup>

#### TRACEABLE PSEUDONYMITY: HOW CAN IT BE DONE?

In today’s technological environment, there are several ways to facilitate a traceable pseudonymous society. In this Essay, I am not confining my analysis to tools that only facilitate pseudonymous “speech,” but also address applications that support a variety of actions, interactions and transactions —in other words, most mundane activities. One “low-tech” option for facilitating pseudonymity is to provide individuals with several Social Security numbers, while only specific organizations,

---

166. For example, in *McVeigh v. Cohen*, 983 F. Supp. 215 (D.D.C. 1998), *supra* note 138, the main issue concerned the disclosure AOL made to the U.S. Navy about the real personality behind a virtual persona.

such as the IRS and the FBI, hold the index to all the personalities and are able to link the virtual and physical entities. Note that this solution does not in fact create virtual personas that cannot be linked to the individual, but splits the individual's identity into several sub-personalities.<sup>167</sup> The use of multiple credit cards, E-Z passes, or even phone numbers and email addresses can facilitate such pseudonymity as well.<sup>168</sup> In the Internet setting, possible solutions may include allocating several usernames to every physical user by an Internet Service Provider or other intermediary. Every name will leave an e-trail of the user's Internet activity, but this trail will be distinctive from the individual's other usernames and user trails.<sup>169</sup> The government can promote the use of pseudonyms in several ways. For instance, the federal tax system can make expenses related to the obtaining of several pseudonyms (e.g., increased payments to ISPs) tax deductible. Government can also act directly through public relations campaigns, which will explain how pseudonyms could be obtained and used, and demonstrate how such uses can mitigate privacy concerns. An important factor to consider when constructing a pseudonymity-based scheme is the number of pseudonyms every user should be allotted or allowed to use. Providing only a few virtual personalities will render the solution ineffective, as each personality will contain sufficient information for the privacy concerns to reemerge. This will occur because a profile of each pseudonym may include enough data to allow problematic practices such as discrimination and manipulation.<sup>170</sup> In addition, with only a few pseudonyms allowed, the costs of abandoning or switching pseudonyms will increase. However, allowing individuals to use an unlimited, or very large, number of pseudonyms brings us back to the anonymous society, with its lack of accountability and other social problems. The tension between having too many and too few identities is evident throughout the following analysis, which provides no clear answer as to what this number should be. Only future research and analysis can reveal the correct

---

167. This solution will cause governmental databases to multiply in size, but recent advances in storage and analysis of data ought to enable such entities to overcome this difficulty.

168. Clearly such pseudonymity schemes will prove effective only if a separation between the various forms of identity could be maintained – an issue I address below (pt. III.4(a)).

169. AOL provides its members with several usernames for browsing online. However, for these tools to be effective, a pseudonymity scheme must ensure that the host websites cannot link the virtual personalities through the use of cookies or IP addresses. This issue is directly addressed in JEFFREY ROSEN, *THE UNWANTED GAZE* 175 (2000). Here Rosen addresses tools developed by Zero-Knowledge.com that allow users to have five digital pseudonyms, as well as the multiple usernames AOL provides.

170. These problems may arise in every instance in which users make extensive use of one identity — be it real or an alias. For an analysis of this issue in virtual worlds, see Tal Zarsky, *Information Privacy in Virtual Worlds – Identifying Unique Concerns*, N.Y.L. SCH. L. REV. (forthcoming 2004) (on file with author).

answer.<sup>171</sup>

Today, many users are already employing multiple email addresses and usernames. This is indeed a promising start. The pseudonymous society I am addressing, however, requires structural and regulatory changes relating to the traceability of the pseudonym, the ways in which it could be used and the number of pseudonyms allotted to each user.

#### PSEUDONYMITY AND IDENTITY

The pseudonymity that the aforementioned applications provide can be understood as a layered protection of identity. In a way, a pseudonymous society calls for constructing two forms of virtual walls: (1) those blocking the connection between the physical identity and the aliases; and (2) those blocking the connection between two or more pseudonyms that belong to the same physical person. The ability to keep these walls intact is interconnected: Once it becomes clear that pseudonym *A* actually refers to person *X* and that pseudonym *B* belongs to person *X* as well, we can immediately deduce that pseudonyms *A* and *B* stem from the same persona. However, the exposure of the connection of only one pseudonymous identity to the physical self need not jeopardize the use of the other identities that can still remain unconnectable to the physical persona. Of these two forms of walls, it is the wall between the real and virtual persona that is intuitively considered to be of importance, as it will allow individuals to conceal their identity, facilitate intimacy, and protect from embarrassment. However, the second type of wall is of importance as well, as many of the problems created in today's information environment do not stem exclusively from the link between our actions and our one physical identity. In a world where the pseudonym is constantly used and surveyed, the privacy of the pseudonym itself becomes a concern.<sup>172</sup> For example, for entities practicing price discrimination or attempting to tread on our autonomy through the use of manipulation, there is no real importance as to who we really are in the physical world. The success of such schemes depends on the extent of knowledge that content providers can obtain about our habits, preferences, and patterns of behavior, and whether they can successfully use this knowledge to their benefit. Thus, returning to the previous example, the fact that pseudonym *A* is actually *Mr. X* is interesting and relevant to information collectors only if they can accumulate information about *Mr. X* that they can incorporate into the profile they are constructing and therefore can apply to pseudonym *A* as well, and vice versa. Therefore, in a world where individuals are all interacting while using pseudonyms

---

171. My intuition is that the use of four or five personalities will lead to an optimal result.

172. See *supra* note 170.

(such as certain parts of the web), revealing the real identity will not be as important as revealing the other pseudonyms *Mr. X* uses. Data collectors will be very interested in revealing these other pseudonyms so that they can consolidate the information collected about different aliases, as well as coordinate the feedback they are sending pseudonyms stemming from the same actual person. Thus, to fully benefit from pseudonymity, it is important to not only use an alias detached from the physical self, but also to use several aliases. If the walls between a user's aliases shatter, the problems of price discrimination, the autonomy trap, and even abuse and misuse will reappear in a pseudonymous society, even though the walls protecting the real identity of the individual are safely in place.

However, it should be noted that in many cases the wall between the aliases and the physical identity is one of extreme importance, and applying information about the alias back to the physical identity could cause substantial detriments. Salient examples are actuary calculations for setting insurance rates, credit ratings, and interest rates.<sup>173</sup> In such cases, any additional information about the alias, such as financial activities or even sport preferences, can adversely affect the actuary result of credit ratings or insurance rates that pertains mostly to the physical identity. Although these are the examples that are usually referred to in discussions of the impact of the lack of information privacy,<sup>174</sup> a pseudonymity-based solution should strive to address all the problems of today's information society — including those of discrimination and manipulation, which are not as obvious and intuitive. An overall solution will be possible only if both walls addressed above remain intact.

### III.2. *Pseudonymity, the Data Flow, and the Detriments of Surveillance*

To understand the effects of traceable pseudonyms and the way they solve the problems of information privacy, I return to the information flow paradigm of the collection-analysis-use of personal data. As with anonymity, regulation that facilitates a shift toward the use of several pseudonyms does not impede the collection or analysis of personal data, nor does it regulate the ability to use the results of such analysis, as the actual change is not in the gathering but in the quality of the data collected.<sup>175</sup> Of the three stages of the information flow, it is the second

---

173. For a discussion of how the issues of medical insurance and credit rating require specific analyses of information privacy concerns, see Zarsky, *supra* note 1, pt. IV.

174. *Id.* See also Garfinkel, *supra* note 80, at 25 (regarding credit ratings); Smith, *supra* note 40, at 312 (same).

175. This is due to the fact that it cannot be linked to a physical person or other virtual personas.

(analysis) where pseudonymity has severe repercussions. As a result of a shift to a pseudonymous society, the analysts cannot connect the information they gathered to any physical persona, but only to a recurring virtual one. In addition, the collectors' databases will include fewer variables about each individual they are profiling. Finally, the number of profiles that the collectors must construct will multiply in step with the number of virtual personalities every physical person uses. These changes will encumber the data analysis and mining process, and in many cases will prevent analysts from discovering meaningful patterns in the personal information databases. To simplify this point, I present the following example:

*Before the shift to the pseudonymous environment, Database A collected 10,000 snippets of personal information a month from various sources, such as websites and catalog marketers. These bits of data were classified, on average, as 100 different variables descriptive of 100 different individuals. The operators of Database A used this information to form consumer profiles of these 100 individuals.*

*Upon shifting to a pseudonymous society, Database A will still gather 10,000 new bits of data a month. However, the internal classification of these bits of data will drastically change, as the database will now include 500 identities with only 20 bits of data pertaining to each identity. This change could have been very good news for the collectors had it resulted from a 500% growth in clientele, but that is not the case — it results from the fact that every individual is now using five alternative personalities. With this new database, the collectors will have less insight into the individuals' profiles and personal preferences, and cannot construct sophisticated patterns to predict future behavior.*

Moreover, the final stage of the information flow paradigm (use) will be affected by the shift to pseudonymity, thus diminishing several concerns stemming from today's information environment.<sup>176</sup> With pseudonymity, there will be fewer opportunities to abuse personal information, as a user can control the availability of potentially abusive information by ensuring that such information can be linked only to one dimension of his or her persona. If an individual uses a virtual persona whose reputation the individual does not value, then he or she will not fear that sensitive information would leak to other pseudonyms whose reputation is more highly valued. In addition, with pseudonymity, concerns about the use of personal information for discrimination and manipulation should diminish as well. As I explained above, informa-

---

176. In discussing the various detriments stemming from the implementation of personal information, I use the same framework as above in addressing fears of misuse, discrimination, and manipulation (*see* pt. I.2, *supra*).

tion collectors require a great deal of information about every user to successfully apply any of these problematic schemes. However, with only partial profiles in hand, marketers and content providers will lose much of their ability to correctly analyze personal information and predict the user's vulnerabilities for their exploitation in terms of discrimination and manipulation.

When examining a potential shift to pseudonymity in the online realm, all the mentioned elements apply, with one important addition. In this context, pseudonymity impedes the ability of marketers and other information providers to benefit from feedback loops they can form upon interacting with their users.<sup>177</sup> As mentioned, today's Internet environment allows information providers to monitor the actions of their users and react by providing them with specific content. However, the data flow does not end at this point, as providers closely examine how the information they provided affects the users' subsequent online behavior, and the providers continue to adjust their responses in accordance to those reactions.<sup>178</sup> Scholars have recently noted that this feedback loop may enable information providers to mold the future responses they receive and even manipulate their users by means of specific content.<sup>179</sup>

However, by adopting pseudonymity online, concerns about this feedback loop and its consequences should be substantially mitigated. In a pseudonymous society, the feedback that information providers receive from users as part of this feedback loop is considerably weaker and less informative, as it refers to only one pseudonym out of a possible many, and therefore only a part of a user's overall behavior and identity. Moreover, every user is not in one feedback cycle, but will be subject to several feedback loops, depending on the number of aliases he or she uses. Every one of the user's respective pseudonyms might be provided with specifically tailored content, with a specific content provider attempting to predict the user's response. But every one of these content providers, and the feedback loops they create, will be pushing different forms of information by using a different profile, which, if used wisely, could portray different aspects of the users' life and activities. Thus, the individual will receive content from diverse sources that will potentially

---

177. I explain the significance of the potential feedback loop that content providers can create in the online environment above, when addressing the fears of manipulation in a transparent society (*see* pt. I.2(d), *supra*). For an additional analysis of the potential detriments of the feedback loop dynamic online, *see* Zarsky, *supra* note 3, at 43.

178. *Id.*

179. *See, e.g.*, LESSIG, *supra* note 79, at 153. This issue is also addressed by OSCAR GANDY, *THE PANOPTIC SORT* 230 (1993).

offset attempts to discriminate or manipulate.<sup>180</sup>

Pseudonymity will prove valuable not only by creating obstacles in the information flow, but by directly empowering users. However, unlike transparency, which empowers users by providing them with additional information about others, or anonymity, which frees the individual from any accountability, pseudonymity-based schemes empower users by allowing them to switch between personalities and in that way enjoy all that the new information society has to offer.<sup>181</sup>

With pseudonymity, cautious users can control the breadth and content of every one of their pseudonym profiles that data collectors are constructing in the background. The key to success in this task is remaining conscious of the specific information submitted and the actions carried out while using every persona. When bearing these factors in mind, an occasional shift between personalities can turn some of the privacy concerns into advantages. Instead of being subjected to discriminatory practices, pseudonym users will receive several offers and prices to their various personas, allowing them to choose between these offers. Rather than subjecting themselves to various forms of manipulation, users can create different personalities, perhaps reflecting the different ideologies and beliefs they are currently contemplating, and see what information and responses they receive through the feedback cycles described above. Users can start every session of shopping, browsing or general exploration in the Internet environment by asking themselves, "Who am I this time?"<sup>182</sup> A user can then pick a personality from his or her bag, and see what responses the chosen persona invokes. When using such aliases, individuals will not fear that collectors might record and use information about their locations, actions, and transactions, but will be counting on it. With pseudonymity, therefore, users can get a taste of different cultures and ideas, while reserving their ability to switch back to their real life unnoticed. Finally, when a user grows tired of a specific virtual personality, or is unhappy with the feedback it generates, he or she can simply set it aside, and move on to explore other opportunities with alternative pseudonyms. With this scheme, users can control their destinies by independently deciding

---

180. As mentioned above with regard to the mitigating effects of transparency, an effective antidote to manipulation is that the individual receives several forms of diverse content (*see supra* notes 107-110 and accompanying text).

181. This is similar to the one mentioned by SHERRY TURKLE, *LIFE ON THE SCREEN* 177 (1995). Turkle argues that the Internet allows users to think of their identity in terms of multiple selves, rather than a single self.

182. Here I am paraphrasing an idea from a short story of this name from KURT VONNEGUT, *WELCOME TO THE MONKEY HOUSE* (1961).

when they wish to start a new chapter in life and turn a page, without fearing that their previous conduct and beliefs will haunt them.

### III.3. *Pseudonymity and the Troubles of Anonymity*

Even though pseudonymity somewhat resembles the anonymous society described above, the harsh criticisms pointed at anonymity are relevant only in part. First, in a pseudonymous society, the fears that infringement, racketeering, and terrorism can be carried out behind a veil of anonymity are somewhat curtailed. In a pseudonymous society, copyright owners and law enforcement can trace problematic and prohibited recurring actions and transactions to particular pseudonyms. Should these tracking entities suspect foul play on the basis of the limited patterns available, courts or other legal authorities can remove either or both of the two walls pseudonymity creates.<sup>183</sup> Pseudonymity's appeal in this context is that it allows for an intermediate state of privacy if only partially incriminating evidence is available — as courts can mandate the removal of only one layer of concealment to allow further investigation, while still maintaining a level of privacy.<sup>184</sup>

Similarly, the problems of accountability, or lack thereof, that dominate any discussion of anonymity are somewhat mitigated in a pseudonymous environment. With pseudonymity, the faces and personas we use and encounter are consistent. Therefore, users have an interest in creating goodwill in their new identities because they use them over an extended period of time, as opposed to the one-time, disposable identity that anonymity provides. As a result, users will refrain from indecency and other problematic forms of behavior when speaking, acting, or interacting through the use of a consistent alias.<sup>185</sup> Regarding these matters, the number of identities each user will be permitted to apply would again prove to be a crucial element that must be considered and regulated, as providing users with many identities will return us to the trou-

---

183. I concede that in the anonymous society government and copyright owners may appeal to courts to lift the veil of anonymity as well. However, pseudonymity still creates an environment that is easier to track for the following reasons: (1) with the use of encryption, re-mailing systems, and other innovations, it will perhaps be impossible to reveal the real identity of the user; and (2) pseudonyms are consistent personalities that are easier to track over time.

184. For example, if the government or even owners of intellectual property are suspicious of the actions of a specific pseudonym, a court might grant them the right to tap into the additional pseudonyms that the individual is using without receiving information linking those pseudonyms to the physical person behind the mask.

185. BRIN, *supra* note 14, at 247, introduces an interesting solution to the problems of accountability that might still persist in a pseudonymous society. Brin suggests that a proper balance between accountability and anonymity might be achieved by requiring all connections between the real and virtual selves to be revealed after the lapse of a designated period of time.

bles of anonymity — the total loss of accountability, and law enforcement's resentment of this solution.

Moreover, unlike anonymity, pseudonymity will still facilitate some analysis of personal information. In a pseudonymous society, collectors can derive beneficial knowledge about users by analyzing the pseudonyms' profile for patterns in ongoing behavior, consumer preferences and trends, while the individuals can somewhat control the amount of information they are submitting by switching between pseudonyms. Therefore, a shift to pseudonymity will not lead to a complete loss of all benefits arising from the analysis of personal information.

Finally, the use of pseudonyms could be simpler than employing anonymizing software. Rather than applying tools of encryption or remailers, individuals will be required only to learn how to use several identities while carrying out their online, and perhaps offline, activities. Tools could be installed in Internet browsers or ISP portals<sup>186</sup> to facilitate this solution and could be set as a simple default.<sup>187</sup> The sophistication challenge that pseudonymity presents is not one of technological knowledge and ability, but of explaining the importance of using pseudonyms in general, the use of several identities in particular, and how shifting among them should be carried out to minimize privacy concerns while maximizing other benefits of the information society.<sup>188</sup>

Of the three global solutions addressed throughout this Essay, pseudonymity provides us with the best and most pragmatic global solution. In a way, a pseudonymous society presents a balance between our desire to use the rich personal information landscape now available, our privacy needs, and the ability of governments to track down lawbreakers. However, before we wholeheartedly advocate a shift to a pseudonymous society, we must acknowledge some of its possible shortcomings.

### III.4. *The Troubles of Pseudonymity*

#### (A) KEEPING UP THE WALLS

Pseudonymity presents an obvious flaw. To maintain a traceable pseudonymous environment, the two walls described above — the walls (1) between the various identities and the physical persona, and (2) among the various identities themselves — must remain intact. The strength of these walls, however, will be constantly tested.

---

186. AOL already allows every user to pick among several identities when logging on.

187. For a discussion of Zero Knowledge technologies that facilitate these services, see *supra* note 169.

188. I concede that the situation changes if the entire net shifts to anonymity, thus creating an environment in which anonymity will be the default reality. In that case, anonymity will not require technical expertise.

At first, a pseudonymous society will rely on trusted intermediaries that connect the aliases to the physical person and transfer information between them. But whom can we trust for this delicate assignment? There is always the fear that those facilitating pseudonymity will sell or abuse the sensitive information they are entrusted with. Therefore, a pseudonymous society must introduce rules prohibiting intermediaries from sharing or using such information. These rules must also specify what legal measures could be taken against those who breach the trust of their customers and reveal these connections; the rules must also specify the instances in which information can be revealed to law enforcement or others.<sup>189</sup>

Those interested in connecting the lines between the virtual and real personas can, however, achieve their objective without obtaining the actual index from its safe keepers, but through data analysis. By reviewing personal data, analysts will try to connect the pseudonyms and the individual's physical identity by using specific forms of personal data that remain constant, such as home addresses, mailing addresses, and phone numbers. This is in fact how credit bureaus are able to construct an overall picture of a consumer's spending even when the individual uses multiple forms of payment. Thus, an overall pseudonymity scheme requires the introduction of new identifying elements that are stable but not easily linked to the physical identity, such as several email addresses or Social Security numbers. With these new identifiers, commercial entities will be able to build a trusted relationship with clients, and even reach them directly in the event of default, but will be unable to easily link the pseudonym and the physical individual behind it.

The other wall — the one among the various pseudonyms an individual is using — is vulnerable as well. For instance, by using clustering and association rules analysis,<sup>190</sup> data mining tools can attempt to group the virtual personalities that contribute to a database. In such analysis, when two or more aliases show a high level of affinity, they will be considered the same person, thus shattering the wall dividing the different personas. This process may sound futuristic and farfetched, but it's hardly beyond the realm of imagination. Here is an example:

*Mary is an enthusiastic online shopper. To protect her privacy, she uses two pseudonyms for her online transactions: MarA and MarB. MarA is the compulsive browser: She searches at many stores, while viewing different products, sales and sizes. But Mary never uses MarA*

---

189. See *supra* note 138 with regard to the debate as to when the veil between the virtual and actual persona can be pierced.

190. See Zarsky, *supra* note 3, pt. I, for explanations and demonstrations as to how such data mining applications can be put to use.

*in finalizing a transaction. That is when she brings in MarB, who enters the website, goes straight to the product and purchases it. Using this technique, Mary believes she enjoys the best of all worlds, as she is viewed as a browsing and selective shopper with one alias, and a decisive, perhaps impulsive, purchaser with the other. Thanks to the use of pseudonyms, Mary has finally found a way to beat the system.*

*Or maybe not? The e-commerce sites Mary is using have decided to try to overcome the problems the use of pseudonyms poses. They begin to analyze their databases in an attempt to group pseudonyms that refer to the same physical individual. During their analysis, they encounter a minor trend — every time MarA logs out, MarB (which originates from the same group of IP addresses) logs in. Moreover, MarB always purchases a product that MarA previously looked at. Additional correlations between these two usernames lead the analysts to believe that they indeed stem from the same person, and from then on, are treated as such.*

This example is simplified, but it demonstrates the opportunities information collectors have in a world of sophisticated analysis and vast databases. Moreover, not all users will reach Mary's level of sophistication and will use their pseudonyms interchangeably. In that case, it will be even easier for analysts to detect which identities originate from the same person.

The above analysis confronts us with the basic flaw of pseudonymity-based solutions: All aliases initially originate from one person, with one mind, and one personality. As we might have only one pattern of behavior, there is a good chance that our pseudonyms will be linked to each other in one way or another. The analysts need only one opportunity to discover the connection between two different faces of the same person, and there is a good chance that they will get it while constantly monitoring the actions of all individuals.<sup>191</sup> Therefore, further research must establish whether the mining and analysis described above can indeed result in piercing the wall that divides the alter egos originating from the same persona.

The final concern with the feasibility of a pseudonymity solution, and the ability to keep the two walls intact, goes to the sophistication these schemes require. Pseudonymity requires all users to master several complex tasks. At first, it requires them to make use of several virtual identities during their daily lives. For this, they will be required

---

191. Clearly the problem of matching several pseudonyms that result from the same identity is a realistic one only when an individual makes use of a limited number of consistent aliases. However, as I explained above (*supra* note 171 and accompanying text), only a solution that provides for a limited number of aliases should be accepted.

to remember several usernames and passwords. In addition, they must be alert not to provide any information, such as home address, that might be easily tied to their physical identity. Finally, they must use these pseudonyms interchangeably while remembering what data trail each pseudonym is creating.

For many users, any one of these requirements may prove impractical. Therefore, user sophistication may in fact become a crucial bottleneck in implementing a broad pseudonymity scheme. Additional research is required as to whether these tasks can be widely mastered and accepted.<sup>192</sup> Perhaps only future generations will be able to cope with these demands — generations that have grown up with the Internet and are therefore comfortable with the use of several identities at the same time.

(B) “VIRTUAL BABIES AND VIRTUAL BATHWATER”<sup>193</sup>

Another inevitable side effect of a shift to pseudonymity will be a decrease in the knowledge that can be derived from the analysis of personal information. In a pseudonymous society, the databases of personal information are very different from those now existing. Databases in the pseudonymous society can accommodate the creation of limited profiles that contain fragments of individual’s preferences and personal information. The analysis of these users’ profiles will still reveal interesting information, but the patterns and clusters will be far less elaborate and interesting than those available in a society where every individual goes by only one name and identity. When shifting to pseudonymity and conducting data mining or other analyses, the patterns that will be lost are those indicating subtle trends with a low confidence factor.<sup>194</sup> These are the patterns of the greatest interest and importance for today’s analysts, as they reveal unapparent trends of customer behavior. In addition, when using pseudonyms, we will lose crossover patterns — those revealing links between preferences in different areas of commerce or behavior.

Again, our quest to protect the public from abuses of personal information will lead to the loss of a valuable resource — one that is too

---

192. Perhaps the use of several passwords could be applied through the use of biometrics, for example by using a different body part for every identity.

193. Here I am paraphrasing from pt. (II)(1) of Zarsky, *supra* note 1, which bears a similar name and discusses in depth the benefits of personal information analysis and how impediments to collection lead to the loss of such benefits.

194. See Zarsky, *supra* note 3, pt. I, for an explanation of the factors of “support” and “confidence” that are used to measure the various patterns and trends emerging in data analysis and data mining.

valuable to give up without serious consideration.<sup>195</sup> Before adopting a pseudonymity-based solution, additional research must precisely determine how the shift to pseudonymity will affect the ability to derive useful knowledge from the personal information databases and what form of knowledge will be lost with this shift.

#### CONCLUSION:

Our long journey through the land of the three global solutions ends with mixed results and many question marks. Our review of transparency revealed an interesting concept with surprising benefits but unsolvable downsides. Anonymity showed the signs of a wonder drug that cures all, but will never be accepted because it poses bitter side effects. Pseudonymity, however, seems to stroll across the golden path between these two concepts, by providing accountability with partial anonymity, and the benefits of information collection with the protection of privacy. This concept will, however, require additional research and analysis before it is accepted as an overall solution.

Even if these solutions will all be rejected, or will not find their way into the hearts of legislators and the public, our journey has not been in vain. Elements of these global solutions should and will be put to use in the information society. When used in specific contexts and to specific extents, they will be extremely helpful and practical. However, when using these global solutions, we must be aware of the possible shortcomings they create and attempt to avoid them.

Finally, this journey forced us to think outside the box of the information flow, and consider ideas that do not strictly fit within this general paradigm. Indeed, as the opening paragraph demonstrates, sometimes only thinking outside the box can lead to a timely resolution of difficult problems.

---

195. On these issues, see Zarsky *supra* note 1, pt. II.