DESPERATELY SEEKING SOLUTIONS: USING IMPLEMENTATION-BASED SOLUTIONS FOR THE TROUBLES OF INFORMATION PRIVACY IN THE AGE OF DATA MINING AND THE INTERNET SOCIETY

Tal Z. Zarsky

INTRODUCTION
I. SOLUTIONS AND THE INFORMATION FLOW
A. Collection
B. Analysis
C. Implementation
D. Summing Up
II. PREFERRING THE REGULATION OF IMPLEMENTATION TO COLLECTION
A. Virtual Babies and Virtual Bathwater
1. Subsidies for Startups and the Importance of Innovation
2. Creating Value
B. The "Personal-Information-Based" Transaction
1. Myopia and Corrective Lenses
2. Collectees vs. Subjects of Manipulation
III. DOUBLECLICK, INC.—A PRACTICAL PERSPECTIVE
A. Collection
B. Analysis
C. Implementation
IV. IMPLEMENTATION-BASED SOLUTIONS
A. General
1. Fear of Abuse and Misuse of Personal Information
2. The Adverse Effects of Errors
3. Seclusion and Solitude
4. Manipulation
B. Price Discrimination
1. General
2. Exceptions
a. Medical and Life Insurance
b. Credit Rating
V. CONCLUSION

DESPERATELY SEEKING SOLUTIONS: USING IMPLEMENTATION-BASED SOLUTIONS FOR THE TROUBLES OF INFORMATION PRIVACY IN THE AGE OF DATA MINING AND THE INTERNET SOCIETY

*Tal Z. Zarsky**

INTRODUCTION

Our personal information is constantly being recorded, stored and analyzed. Commercial entities watch our every action, storing this data and analyzing it in conjunction with information acquired from third parties. These entities use this knowledge to their benefit (and at times, our detriment) by discriminating between various customers on the basis of this personal information. At the same time, in the media market, large conglomerates can now provide specifically tailored content to individual customers on the basis of such data, thus potentially controlling their perspectives and impairing their autonomy. The expanding use of data mining applications, which enable vendors to search for hidden patterns and associations in a fast and efficient manner, only makes matters worse and accommodates the commercial entities in carrying out these activities.

We obviously have a problem. The first step in its solution is to create public awareness of these problematic practices and their possible outcomes.[1] Focusing public opinion on the important issues is only the beginning, however. Public opinion, which is easily influenced and at times wanders to other areas of interest,[2] is insufficient to cope with the complex and elusive problems of information collection and analysis. Regulation must be implemented in order to protect consumers from these detrimental practices.

There is no shortage of proposed solutions to the problems of information privacy. Prominent legal scholars, governmental committees, and even the commercial entities themselves have proposed regulatory schemes to overcome the problems created by the gathering of personal information.[3] In some countries,

* Resident Fellow, Yale Law School. J.S.D. Columbia Law School, L.L.M. Columbia Law School, LL.B/B.A. Hebrew University of Jerusalem. The Author would like to thank Eben Moglen, Lance Liebman, Paul Schwartz, and the members of the 2002 J.S.D Candidate workshop. The Author also thanks Yochai Benkler and Eli Noam for providing additional insight and assistance regarding this paper.

1. See Tal Zarsky, "*Mine Your Own Business!" : Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion*, 5 YALE J. OF L. & TECH. 2002-2003, at *26-7 [hereinafter *MYOB*], for the content of the public campaign that should be launched to promote these interests. See also LAURA J. GURAK, PERSUASION AND PRIVACY IN CYBERSPACE (1997) for the importance of public opinion in this context.

2. *See, e.g.*, T. Loomis, *Privacy Class Action Suits Over the Misuse of Customer Data Have Run Their Course*, THE NEW YORK LAW JOURNAL, July 25, 2002, at 5. This article indicates the shift in public opinion as one of the reasons for the recent sharp decline in cookie-related litigation. *Id.*

3. There have been several initiatives by associations representing commercial entities of various interests. *See, e.g.*, Network Advertising Initiative (NAI), *at* http://www.networkadvertising.org/default.asp (last visited Aug. 20, 2003).

and in certain areas of law, actual legislation is now being enforced with varying levels of success and compliance.[4] In addition, several countries have created executive branches vested with the authority to regulate and supervise the use and collection of personal information.[5] The solutions provided are diverse and range from self-regulation to the implementation of strict regulation and governmental intervention. They draw from various fields of law and at times require changes in the technological infrastructure facilitating the interaction between the individuals and the information collectors. However, the debate as to the ideal solution to these problems continues, while the changing technological landscape makes choosing an effective solution even harder.

In Part I, the Author explores the various solutions currently contemplated for solving the lingering problems of the flow of personal information. The key element integrated into the ongoing discussion of these solutions is the growing use of data mining applications for the analysis of personal information—a phenomenon that provides collectors and users alike with a new perspective to the world of information, but also creates additional problems and opportunities for abuse of such data. The growing use of data mining applications has a profound effect on the correct definition of problems created in the new information environment[6] and naturally will have a similar affect regarding solutions.

In order to fully understand both the problems and the solutions, this Article first draws out a clear scheme of the personal information flow. This three-part process consists of collection, analysis and implementation of data, and is imbedded in the current legal and technological landscape. Understanding this process will serve us in several ways. First, viewing the process as a whole will allow us to locate the stages at which the problems arise, and perhaps deal with them directly. Second, it will allow us to compare and distinguish between various solutions and fully understand what the implementation of today's proposed solutions will entail. It also allows us to easily view the full effect of such contemplated solutions, with regard to both the desired result and the possible side effects. Obtaining such a wide perspective will assist us in deciding whether these effects are indeed inevi-

---

4. The European Union (EU) adopted the European Directive of Data Protection, effective on October 25, 1998. Council Directive 95/46/EC 1995 O.J. (L.281) 31 [hereinafter "EU Directive"]. *See* PETER SWIRE & ROBERT E. LITAN, NONE OF YOUR BUSINESS: WORLD DATA FLOWS, ELECTRONIC COMMERCE, AND THE EUROPEAN PRIVACY DIRECTIVE (1998). *See also* Fred Cate, *Data Protection Law and the European Union's Directive: The Challenge for the United States: The EU Data Protection Directive, Information Privacy, and the Public Interest,* 80 IOWA L. REV. 431. In the United States, the government introduced specific legislation regarding certain issues of information privacy. *See* Daniel J. Solove, *Privacy and Power: Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393, 1430-45 (2001). Examples of such laws include: The Fair Credit Reporting Act (FCRA) of 1970; The Cable Communications Policy Act (CCPA) of 1984; The Video Privacy Protection Act (VPPA) of 1988; The Driver's Privacy Protection Act (DPPA) of 1994 ; The Health Insurance Portability and Accountability Act (HIPAA) of 1996; The Children's Online Privacy Protection Act (COPPA) of 1998; and The Gramm-Leach-Bliley Act (GLB) of 1999. *Id.* at 1440-44. For a worldwide description of privacy-related laws, see Mike Hatch, *Electronic Commerce in the 21st Century*, WM. MITCHELL L. REV 1457, 1468-70 (2001).

5. For example, Canada created an "Information and Privacy" Commission. *See* Information and Privacy Commissioner, *at* http://www.ipc.on.ca (last visited Aug. 20, 2003). See J. Reidenberg, *Resolving Conflicting International Data Privacy Rules in Cyberspace*, 52 STAN. L. REV. 1315, 1347-48 (2000), for the role of such agencies in the EU.

6. *See MYOB*, *supra* note 1, at *19.

table, or whether other changes and solutions would allow us to avoid or mitigate side effects.

Therefore, using the information flow paradigm, the Author sets out to find the best solution to the problem of information privacy. In order to find a suitable solution, or when forced to choose between several proposed solutions, we must make several inquires. Obviously, we must inquire whether any proposed legislation will solve the problems at hand and eliminate the negative results of today's constant surveillance and information collection. We must further determine the side effects of such regulation and whether these can be avoided by taking an alternative route. In addition, we must inquire whether the outcome of the proposed solution is fair and achieves equity among the participants in the information market. Finally, given the fact that powerful interest groups have high stakes in these issues,[7] solutions must be pragmatic in their requirements of the government and legislators. In this context, governments must face the difficult task of balancing the public demand for protection in the personal information arena against the interests of powerful commercial entities that are interested in using such data for their own benefit.

In Part II, this Article provides several critiques to the regulatory solutions discussed in Part I, which will reconsider the solutions' applicability and perhaps send us searching for alternate regulatory schemes. Such critiques, based on the growing use of data mining applications, conclude that of the three stages of the information flow, solutions should be incorporated into the final implementation stage. This Article further asserts that most of today's proposed solutions do not provide an optimal response to the problems at hand. In many ways, choosing among these solutions resembles selecting a remedy for the common cold—there are many products and brand names to choose from, yet their effectiveness is questionable. Some remedies prove to have problematic side effects and many deal with the symptoms only and not the underlying problems. However, unlike the common cold, the problem of personal information flow will not disappear on its own after seven days.

To further our understanding of the information flow process, as well as the benefits and detriments of the various solutions, Part III provides a practical perspective by analyzing a recent settlement agreement between DoubleClick, Inc. and a group of State Attorneys General. The Article addresses the practices of DoubleClick and the rules and regulations to which it has been subject to in this agreement, and analyzes whether they are appropriate to solve the problems the new information environment creates. This analysis makes use of the information flow paradigm introduced in Part II, and incorporates the various solutions the agreement mandates within this process.

The Article concludes in Part IV by addressing the problems dominating today's privacy debate and providing appropriate solutions in view of the critiques set forth. The Article addresses several solutions generally, and goes on to provide an in-depth solution to the problems of price discrimination. Part IV introduces a

7. The groups with an obvious stake in this matter include information brokers, such as the Direct Mailing Association. *See, e.g.*, Direct Mailing Association, *at* http://www.the-dma.org (last visited Aug. 20, 2003). In addition, various Internet companies and content providers would be interested in preserving their ability to collect information and use it as they deem fit.

new solution that is based on changes in the Internet infrastructure, rather than simplistic regulation prohibiting such practices. To the extent possible, the solutions discussed will be rooted in, and focused on, the final stage of the information flow—the implementation stage. This stage should be regarded as the source of the problem and is therefore the key to any solution.

I. SOLUTIONS AND THE INFORMATION FLOW

Let us begin our journey by generally describing the three-part process of the information flow. A bird's-eye view of the use of personal information helps us to understand the entire process and establish where solutions to the problems addressed elsewhere are best positioned.[8] When taking this broad perspective, the entire flow of personal information can be divided into three distinguishable segments: (1) the collection of personal information regarding relevant individuals; (2) the analysis of the information collected and (3) the use and implementation of the analyzed information for the benefit of the database holders.[9] The Author takes a closer look at these stages, while focusing both on recent technological developments that are having a profound effect on the entire process and on various proposed solutions embedded within every part.

A. Collection

The gathering of personal data in a variety of settings through various means is the first essential stage in the flow of information. This information is later used for various tasks of analysis and implementation. The practice of information collection is far from novel, yet it has been recently enhanced due to several technological and social changes.

First, there has been great progress in the development of surveillance technology. Currently, sensitive cameras as well as powerful wiretapping devices enable the collection of personal information from a distance, without the subject's knowledge of such surveillance and even though the subject's actions have been well concealed.[10] In addition, the use of "simple" means of surveillance (especially cameras) has multiplied due to their diminishing prices, and they are now installed in almost every public area, often at a concealed location.[11] It is unclear what the future has in store regarding this issue, but some science fiction writers

8. *See generally MYOB*, *supra* note 1. In this Article, the Author addressed several problems under the following headings: Discrimination, Fears of Manipulation and Threats to Autonomy, Misuse & Abuse, Intrusion on Seclusion and the Fear of Errors.

9. The use of this three-part structure is not new and variations have been addressed in several articles. *See, e.g.*, Seth Safier, *Between Big Brother and the Bottom Line: Privacy in Cyberspace*, 5 VA. J.L. & TECH. 6 (2000). However, this taxonomy facilitates the critiques provided at a later stage and therefore is extremely helpful. It also allows for correct assessment of the overall effects of every solution.

10. *See* note 29 regarding thermal vision. For a lengthy description of privacy-destroying technologies, see A. Michael Froomkin, *The Death of Privacy?*, 52 STAN. L. REV. 1461, 1468 (2000). In addition, see REG. WHITAKER, END OF PRIVACY, 80 (1999) for more information on surveillance technologies.

11. On the uses of cameras in a wide array of surveillance situations, see Jeffrey Rosen, *A Watchful State*, N.Y. TIMES, Oct. 7, 2001, 36, at 38. *See also*, Froomkin, *supra* note 10, at 1476-79.

12. DAVID BRIN, THE TRANSPARENT SOCIETY: WILL TECHNOLOGY FORCE US TO CHOOSE BETWEEN PRIVACY AND FREEDOM? (1998). In addition, see Froomkin, *supra* note 10, at 1501 (referring to NEAL STEPHENSON, THE DIAMOND AGE (1995)).

describe future "bug cameras,"[12] with the ability to enter our homes without our knowledge and transmit live pictures to their operators. The full computerization of the most mundane activities also enables such in-depth surveillance.[13]

Furthermore, many of our actions and transactions leave a traceable trail behind. Among the reasons for such growing traceability is the extensive use of credit rather than cash (as it is in many cases more convenient), which enables most transactions to be traced back to the individual. In credit card transactions, the collectors are not only provided with information pertaining to the amount paid but also with data regarding the time and place of payment, which may prove extremely valuable during future analysis. Information concerning our location is also available to our cell phone operators, who are able to pinpoint the phone's location according to "messages" our phones continuously send to the nearest antenna.[14] Future use of biometric means of identification will create a society in which almost all actions are traceable back to a specific individual.[15]

The rise of the Internet presents additional dimensions to the collection process. At present, most activities carried out on the web can be easily recorded for later analysis.[16] Moreover, surveillance in the Internet era brings a shift in the sphere of collected information from the "buying client" to the "browsing client."[17] This shift occurs as in the virtual mall; every step a customer takes is watched and recorded. However, growing use of surveillance cameras coupled with the recent progress in facial recognition technology[18] may soon bring these abilities and practices of constant tracking to the brick and mortar city, street and mall as well.[19] In the near future it is quite possible that all our actions will be recorded, and information regarding our locations and whereabouts will be constantly collected by private entrepreneurs eager to use this information to their benefit. Such efforts are supplemented by recent improvements in computer hardware that enable the

13. For example, while in the past, one would pay for a subway ride with a token, now a Metrocard is used, which enables the Metropolitan Transit Authority to track the activities of various users over time. See SIMSON GARFINKEL, DATABASE NATION: THE DEATH OF PRIVACY IN THE 21ST CENTURY 82-84 (2000) for a discussion regarding the use of Metrocards for surveillance.

14. *See* Froomkin, *supra* note 10, at 1479.

15. Froomkin, *supra* note 10, at 1494.

16. The Internet provides a "quantity leap" in the collection of personal data. *See*, *Resolving Conflicting International Data Privacy Rules in Cyberspace*, *supra* note 5. In the Internet setting, surveillance is generally enabled through the use of cookies; however, a growing number of websites are now requiring that the user log on before receiving the site's services, thus facilitating additional collection. For a general discussion of cookies, see *MYOB*, *supra* note 1, at *16.

17. These points are emphasized by Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1198 (1998). In addition, see Paul M. Schwartz, *Privacy and Democracy in Cyberspace*, 52 VAND. L. REV. 1607, 1644 (1999), for a discussion of the extensive data trail and data shadow created in the Internet setting.

18. For a discussion of the uses of facial recognition, see LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 145-46 (1999). *See also*, Froomkin, *supra* note 10, at 1475-79.

19. *See* Stephanie Simon, *Shopping with Big Brother*, LOS ANGELES TIMES, May 1, 2002, at 1 (*referenced in* William Safire, *The Intrusion Explosion*, N.Y. TIMES, May 2, 2002 at 27). This article describes a real store that is used as a laboratory for marketing research. In this store, cameras track the movements of shoppers throughout the store and record the amount of time that they spend at each product showcase. Thus, the Internet's ability of meticulous collection is brought to the physical world.

inexpensive storage of nearly infinite amounts of data. That way, all collected data could be saved and nothing will be lost.

The collectors' right to gather personal information pertaining to others has always been at odds with the individual's right of privacy.[20] This right is protected through the privacy torts created by common law and accepted, to a certain degree, by most states.[21] The realm of the right of privacy, as well as the meaning of privacy in general, has always been hard to define.[22] Yet the realm of protection the privacy torts provide is usually set after taking into account the individual's objective and subjective expectation of privacy.[23] In many cases, both factors are based on the distinction between the public and private domain.[24] Yet in today's age of technology and convergence, the distinction between the public and private domain is not one that is easy to make. It is relatively simple to decide whether you are intruding on the land of another, but somewhat more difficult to determine the legal boundaries of a website you are "visiting" from your computer.

Generally, in the public domain, trespassing individuals assume the risk of being seen, while others maintain the right to gather information. The private domain, however, is usually off limits for such collection. This rule of thumb obviously has exceptions; some courts have found the collection of information in the public domain inappropriate. Such exceptions were usually based on outrageous conduct of the collectors or embarrassing photographs, and in general are a rarity.[25] Thus, the registration of cars passing through a public highway or the positioning of cameras to film public areas[26] presumably do not face any legal restrictions in view of the existing legal doctrines. These legal rules, however, may be subject to change due to a shift in the public's expectation of privacy. Such a shift is quite plausible, as today's (and especially tomorrow's) public areas provide only a false sense of anonymity. However, it is quite possible that the notion of losing oneself among the city masses is gone forever.[27]

A different sort of interaction is the collection of information within the con-

---

20. For a brief discussion of these issues, see *MYOB*, *supra* note 1, nn.9-11.

21. *See* Diane L. Zimmerman, *Requiem for a Heavyweight: A Farewell to Warren and Brandeis's Privacy Tort*, 68 CORNELL L. REV. 291 app. at 365-67 (1982).

22. For the many meanings of privacy in this context, see Daniel J. Solove, *Conceptualizing Privacy*, 90 CAL. L. REV. 1087 (2002).

23. Of the possible privacy torts, intrusion is the most relevant to the collection of personal information. The RESTATEMENT (SECOND) OF TORTS: INTRUSION UPON SECLUSION § 652B (1977) mentions that liability will arise should the "intrusion be highly offensive to a reasonable person"—an objective criteria.

24. Therefore, the property criterion will indicate whether there was a breach of privacy. Prosser takes this opinion to its extreme in stating: "On the public street, or in any other public space, the plaintiff has no right to be alone." William L. Prosser, *Privacy* 48 CAL. L. REV. 383, 391 (1960). *See also*, June Mary Z. Makdisi, *Genetic Privacy: New Intrusion a New Tort?,* 34 CREIGHTON L. REV. 965, 996 (2001).

25. For example, in Nader v. Gen. Motors Corp., 255 N.E.2d 765, 771 (N.Y. 1970), Ralph Nader was followed so closely by private investigators that they were able to see the denominations of the bills Nader withdrew from his ATM account. *See also*, Gallella v. Onassis, 487 F.2d 986, 992 (2d Cir. 1973), in which the paparazzi continuously jumped at Jacqueline Onassis and endangered her children. For a discussion of the importance of the public place factor in privacy cases, see Makdisi, *supra* note 24, at 999.

26. *See* Rosen, *supra* note 11.

27. These possible shifts in the public's expectation of privacy are beyond the scope of this Article and will not be discussed.

fines of the private domain. Here, the law has provided several forms of protection. First, constitutional law protects individuals from the prowling eye of the state within the confines of their home or where individuals have a reasonable expectation of privacy.[28] The boundaries of such protection were recently tested in *Kyllo v. United States*.[29] In *Kyllo*, the Court held that the use of thermal imaging heat detectors by the police to view inside homes in order to detect marijuana-growing lights was unconstitutional. The Court stated that individuals are entitled to retreat to their homes for freedom, and that exploring the inner parts of the home using such a device (that is not in general use) should not be carried out without a warrant. However, this Article will focus on the collection efforts of non-governmental entities. In this area of the law, the use of torts such as intrusion or trespass, provides individuals with protection from the prowling eyes of others.[30] Therefore, the use of many of the futuristic applications mentioned above, such as "bug cameras" or sophisticated tapping devices, will fall within the protection of these torts. Although the use of such devices within the individual's private domain may not be widely practiced, there still are other reasons for concern.

Surprisingly, most of the collection of personal information occurs in an area that is usually beyond legal debate—within the collectors' property, and when gathering information regarding the collector's clients, employees and customers. Those individuals subject to collection, such as shoppers identifying themselves through the use of a supermarket club membership card, may be irritated and frustrated by the surveillance to which they are subject but have a very weak claim of breach of their perceived privacy. This is because their entrance into the collector's domain (both on and off line) could be construed as implied consent to any form of surveillance or information collection.

The Internet environment presents many opportunities for collection within the collector's domain. For instance, website operators may collect data regarding the website user's click pattern, browsing method, and time of entry and purchase. All such information pertains to actions that take place within the confines of the website and are usually collected lawfully.

In addition to the problems discussed above, the use of tort law in this context requires us to assess the damage that surveillance causes. This can prove to be a difficult task that requires us to confront the subjective views of the various individuals subject to such surveillance.[31] This yet again exposes the vulnerability of tort-based solutions to the troubles of online and offline privacy.

In view of these difficulties, many scholars have concluded that the current privacy torts and rhetoric are mostly unsuitable to deal with the problems arising from the collection of personal information.[32] Therefore, an alternative perspec-

---

28. For a brief discussion of the Fourth Amendment in this context, see Fred H. Cate & Robert Litan, *Constitutional Issues in Information Privacy*, 9 MICH. TELECOMM. & TECH. L. REV. 35, 42 (2002).

29. 533 U.S. 27 (2001).

30. *See* Makdisi, *supra* note 24, at 984.

31. In the Internet surveillance context, the inability to prove damages was one of the reasons for the constant failure of cases advocating against the installment of cookies. *See MYOB*, *supra*, note 1, at *15-17. *See also* Froomkin, *supra* note 10, at 1528 (stating that the difficulty in proving damages deterred class actions for these issues).

32. *See* Kang, *supra* note 17, at 159 (referring to various scholars' belief that information privacy can not be solved by the use of today's privacy torts).

tive to the collection paradigm is now gaining popularity—one that views the information collection process as a *transaction* between the collector and the collectee[33] rather than the collection of information from a passive subject. In this transaction, the collectees are conceding or "selling" their personal information by allowing others access to such data[34] or by actively providing it in exchange for convenience or actual compensation.[35]

When taking this perspective, our days are filled with such information-based transactions. For instance, in deciding to use many of the means of modern convenience we are actually forming a contract premised on our personal information. When preferring the use of credit to cash (leaving a transactional trail behind us), cell phones to pay phones (allowing the phone operators to pinpoint our exact location), or even the virtual store as opposed to the physical one, we are partaking in such transactions.[36] Future uses of biometrics as means of identification, which are now in development, add an additional layer to these personal information-based transactions; in the near future, the use of our thumbprint or retina will allow us to leave our wallets at home, but create an information trail following us wherever we go.[37]

The collection of personal information can be viewed both as a potential tort and as the outcome of an information-based transaction. These two perspectives view the same event in two different ways. Drivers passing through a tollbooth using an E-Z pass payment system can be viewed as being subject to the collection of personal information. On the other hand, these actions could be viewed as a transaction between the drivers and the E-Z pass operators. The drivers are receiving services, convenience, and at times, reduced rates for the toll, while providing payment and personal information. Such a "double perspective" could also be applied to view the use of supermarket club cards in purchases, cookie-enabled Internet browsers, and a non-caller ID blocker. Taken to its extreme, the surveillance of an individual walking the streets and into a store could be viewed as either an "active-passive" interaction, or a form of a social transaction (as the individual could decide to wear a mask or simply not walk outside if she preferred not to be

---

33. "Collectees" is not a word. However, the Author uses the analogy of "tipper" - "tippee" to define a new word that will describe the person whose pertinent information is being collected. Use of words like "subject to collection" are inappropriate, as they lack the neutrality needed at this point. In the EU Directive, the terms used are "subject" and "controller." These terms are also improper because they allude to the imbalance of power between the two parties. *See* SWIRE & LITAN, *supra* note 4.

34. Froomkin addresses a dichotomy between transactional and surveillance information. Froomkin, *supra* note 10. The Author disagrees with the way this dichotomy is presented, as many of the forms of surveillance could be viewed as transactions and vice versa.

35. For defining the collection of personal information as a transaction between two parties, see Richard Murphy, *Property Rights in Personal Information: An Economic Defense of Privacy*, 84 GEO. L.J. 2381, 2402 (1996).

36. *See* Froomkin, *supra* note 10, at 1479.

37. Even without the use of biometrics, such practices are now a reality through the use of "SmartCards" for a growing variety of transactions. *See* CNN.com *Japan Seeks Smarter Ideas for Smart Cards*, (Feb. 17, 2003) *available at* www.cnn.com/2003/TECH/ptech/02/17/japan.smart.cards.ap/.

38. Froomkin, *supra* note 10 at 1532-33. Froomkin suggests that perhaps people should begin to wear masks. *Id.* at 1532. Froomkin further states that perhaps at this time, laws that forbid wearing masks should be changed. Such laws exist in several southern states and were enacted as a reaction to the Ku Klux Klan. *Id.* at 1532-33.

seen).[38]

Various forms of solutions and regulations rooted in this stage of the information process[39] have been proposed, and to some extent, already implemented. The prohibition on the collection of information seems to align with preexisting moral and legal doctrines, which frown upon "Peeping Toms" and protect people's faces and names from being used for financial gain without their approval.[40] Solutions based in this phase also align with now mythological fears of "Big Brother" seeing all and gathering information.[41] Therefore, providing solutions to the problems created due to the accumulation of personal information by regulating the collection process seems to be the proper response. However, after additional scrutiny, this initial instinct proves to be wrong.

In general, tort-oriented solutions suggest widening the realm of privacy torts protection from collection beyond the collectee's home and into the public arena.[42] This expansion is justified by recent changes, which subject almost all individuals to the possibility of constant surveillance outside their home, where each of their actions can be seen and recorded. Such an extension of the zone of privacy can be achieved by creating new torts, which will enjoin entities from the collection of personal information or mandate restitution when such collection takes place. This objective could also be met by implementing rules that forbid the collection of specific forms of information or the installment of surveillance tools in certain areas.[43] Yet expanding the zone of privacy in this way might be difficult, as in many cases the collectees provide their consent to the collectors' actions (implicitly or explicitly).[44] In addition, the over-expansion of tort protection from invasive behavior and surveillance to include actions carried out publicly directly clashes with the collectors' First Amendment rights, as free speech must include the freedom to listen and collect information.[45] In view of these difficulties, several solu-

39. Many of such solutions no doubt pertain to the later stages of the information process and regulate the way information should be analyzed, sold and used. However, this article posits such solutions in the collection stage, as it is at this point that the terms of future use are decided upon.

40. The tort of "misappropriation" (that protects a person's face and likeness) has been created by common law. However, it has been backed by legislation in several states. *See, e.g.,* N.Y. CIV. RIGHTS §§ 50, 51 (McKinney 1903). This law was introduced after this right had been denied in New York. Roberson v. Rochester Folding Box Co., 64 N.E. 442, 443-44 (N.Y. 1902). *See also* George P. Smith, II, *The Extent of Protection of The Individual's Personality Against Commercial Use: Toward A New Property Right*, 54 S.C. L. REV. 1 (2002).

41. On the fear of "Big Brother," see generally Solove, *supra* note 4. An additional metaphor of constant surveillance used frequently is Bentham's "Panopticon"—an ideal prison that allows the guards to observe all the prisoners at all times. *See* Whitaker, *supra* note 10 at 32-33.

42. For an example of scholarly suggestions to enact additional torts, see Jessica Litman, *Information Privacy/Information Property*, 52 STAN. L. REV. 1283, 1313. *See also* Jeff Sovern, *Protecting Privacy With Deceptive Trade Practices Legislation*, 69 FORDHAM L. REV. 1305, n.57 (detailing various ways to use the torts paradigm to overcome the troubles of information privacy).

43. For example, the Cable Communication Policy Act (CCPA) prohibits the collection of information that is not necessary for cable service. 47 U.S.C. § 551(b). Collection restrictions are part of CCPA as well. *See* Solove, *supra* note 4, at 1443.

44. Litman discusses this issue and suggests that only a specific form of informed consent (as used in medical law issues) should be accepted. *See* Litman, *supra* note 42, at 1311.

45. For a recent analysis of this issue, see generally Cate & Litan, *supra* note 28.

tions have emerged, which focus on the transactional aspects of the collection phase.[46] An interesting approach, famously advocated by Lawrence Lessig, attempts to solve the problems addressed above by technical means and proposes that solutions are best implemented in the infrastructure, or "code" of the relevant systems that facilitate the transaction.[47] Using the Internet as a prototype for these solutions, Lessig and others suggest that software tools such as P3P assist collectees when "bargaining" with collectors.[48] These applications serve as intermediaries between the collectees and the collectors by automatically sifting through the various information providers and vendors, and only permitting transactions with those who provide for a level of privacy above a certain minimum in their information practices.[49] Such tools indeed show promise, yet require a level of sophistication that the average user does not always maintain.[50] In addition, the ability to transform similar applications to the physical world (as opposed to the Internet) is questionable. In the physical world surveillance continues, yet agents (such as P3P) are not easily incorporated into the "brick and mortar" reality.[51]

Another regulatory scheme that sets out to solve the problems of information privacy and is based in the collection stage is the notice requirement.[52] This proposed solution requires collectors to inform the collectees of the collection process and of the possible uses of the information collected. The notice requirement is considered by many as insufficient, as the collectees tend to be indifferent to such notices, or are incapable of grasping the full impact of the future uses of their personal information.[53] The notice requirement is usually coupled with a right of choice, which will allow the collectees to avoid their inclusion in the collector's database.[54] A great amount of writing and debate focuses on the proper meaning

---

46. A transactional solution is preferred because it could provide one set of rules and norms for those collecting the information, and another for those who might purchase it on the secondary market. *See* discussion *infra* Part I (regarding secondary use and secondary sale).

47. *See* Lessig, *supra* note 18 at 160-61. For a critique of this solution, see Marc Rotenberg, *Fair Information Practice and the Architecture of Privacy*: (*What Larry Doesn't Get*), 2001 STAN. TECH. L. REV. 1, (2001).

48. *See* LESSIG, *supra* note 18 at 160-61.

49. *See* LESSIG, *supra* note 18 at 160-61.

50. Paul Schwartz, *Beyond Lessig's Code for Internet Privacy: Cyberspace Filters, Privacy-Control, and Fair Information Practices*, 2000 WIS. L. REV. 743, 754 (2000), Schwartz addresses these issues as the "blinking twelve" problem, and noting that the "blinking twelve" on most peoples' VCRs results from their inability to program the timer. *Id.*

51. Surveillance is escalating to soon share the heights of the online environment. *See* discussion *supra* note 19.

52. The FTC lists "Notice" as one of the five elements of Fair Information Practices. FED. TRADE COMM'N, PRIVACY ONLINE: FAIR INFORMATION PRACTICES IN THE ELECTRONIC MARKETPLACE: A FEDERAL TRADE COMMISSION REPORT TO CONGRESS (May 2000) (hereinafter Fair Information Practices) *available at* http://www.ftc.gov/os/2000/05/index.htm#22. In its 1998 report, the FTC states that "Notice" has been required in other countries since 1973. *Id.* It is also an element in the NAI principles. NAI, *supra* note 3. For the history and content of the Fair Information Practices, see Schwartz, *supra*, note 50, at 779.

53. For a critique of the notice requirement, see Kent Walker, *The Costs of Privacy*, 25 HARV. J.L. & PUB. POL'Y 87, 106-07. Walker addresses five problems of such notices: (1) overkill (2) irrelevance (3) opacity (4) non-comparability (5) inflexibility. *Id.* at 106-113.

54. "Choice" is one of the "Fair Information Practices" as well. FTC, *supra* note 52.

of "choice," especially regarding the question of whether the choice should be one of "opt in" or "opt out."[55] Yet it is unclear whether there is a practical difference between these two options.[56]

Another solution for regulating the collection of personal information proposes several changes to the contractual default rules governing these information-based transactions.[57] These new defaults (as opposed to the current legal landscape) will assist collectees by initially protecting several specific rights of the individuals. For example, Paul Schwartz, Jeffery Kang and others suggest that these defaults include rules stating that the information submitted must not be sold to another party or shared with affiliates.[58] The shortcomings of such a solution are obvious–-defaults are only defaults. By "lawyering" around these rules, their effect could be easily minimized. To avoid such defaults, sellers will add an agreement to every transaction, in which the collectees will provide consent to the collection and use of their personal information. Thus, the entire effect of this solution would be lost. Still, however, such default rules might assist in bringing the consequences of surveillance to the public's consciousness.

Another interesting and popular solution with regard to these transactions is the use of property rights.[59] This solution attempts to resolve the problems of information privacy by providing the collectees with property rights in their personal information.[60] In view of such rights, a collector will be obligated to license or purchase the information from the collectee as part of the collection process if interested in analyzing this personal data for future uses. Therefore, this solution assures that personal information will not be collected without the consent and perhaps appropriate compensation of the relevant collectee.[61]

The property solution was well accepted at first, but has come under heavy criticism for various reasons. First, critics state that the creation of a property right in personal information will have the opposite effect of the one desired and will actually promote the collection and use of personal information.[62] The "commodification" of personal information will motivate the trading of such data,

---

55. This debate focuses on the question of whether users should actively permit the use of their personal information, or must actively "opt out" of such use. *See* U.S. West v. F.C.C., 182 F.3d 1224 (10th Cir. 1999) (vacating the FCC rule, which mandates the shift to an "opt in" rather than an "opt out" rule in the telecommunications context). *Contra* Hatch, *supra* note 4, at 1495-1501.

56. As explained below, the disparity between the collectors and collectees market power, knowledge, and understanding will render the results of any transaction unfair.

57. Murphy, *supra* note 35, at 2414 (analyzing suggested changes in the default rules).

58. *See generally* Schwartz, *supra* note 50; Schwartz, *supra* note 17; Kang, *supra* note 17, at 1246-48.

59. *See generally* Symposium, *Cyberspace and Privacy: A New Legal Paradigm*, 52 STAN. L. REV. 987 (2000).

60. The concept of privacy as property can be dated back to John Locke and was probably introduced to this context by Alan Westin in the 1960s. FRED H. CATE, PRIVACY IN THE INFORMATION AGE 21-23 (1997).

61. For a discussion as to why property rights are appropriate for this problem, see Julie E. Cohen, *Examined Lives: Information Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1379 (2000). As Cohen states, the leading reason for applying property rules to these problems is that we cannot help but apply a property paradigm to this field of law as well. *Id.*

62. Mark A. Lemly, *Private Property*, 52 STAN. L. REV. 1545, 1551 (2000).

63. Litman, *supra* note 42, at 1290. *See also*, Lemly, *supra* note 62, at 1551 (stating that the "propertization" of personal information will lead to a right that is regularly signed away).

thus increasing the problems we are trying to solve.[63] Other critics point to inherent problems in the data transaction that will block the emergence of a fair market for personal information and create one with considerable advantages for the information collectors. Such problems stem from the collectors' unequal advantage in their understanding of the possible benefits and detriments stemming from the use of personal information and other high transaction costs.[64] A look at the jurisprudential justification for the existence of property rights also leads to the conclusion that the use of a property regime in this context is somewhat problematic.[65]

In addition to gathering information directly from the collectees, today's collectors have other ways to obtain personal data. At first, information is collected from a variety of open sources provided by the government or state agencies. These public records present rich demographic information in various categories, such as the population residing in a specific geographic area.[66] Even though such data might seem harmless at first, as it usually does not refer to specific persons, collectors make use of such information by analyzing it in conjunction with other databases.[67] In view of these possible practices, several commentators note that the government must intentionally tamper with the accuracy of such census information to avoid harmful uses by commercial entities.[68]

One of the greatest controversies regarding data collection pertains to the use of personal data collected by others and acquired on the secondary market. Various collectors, as well as information brokers who gather information and sell it by the "block" facilitate this market.[69] Not always are the secondary sales regarded as part of the collection process, though they are clearly an integral part of structuring the collector's personal information database. In many cases, those acquiring personal data use it for different tasks than those uses for which the data was originally collected; in other words, the data is put to "secondary use."

Several proposed regulatory schemes focus on the problems of this secondary market. Such proposed solutions mandate the consent of the collectees prior to the sale of their personal information—as the collectees are rarely even aware of the secondary sales of their information when they originally provide it. These re-

---

64. *See* discussion *infra* Part II.

65. Cohen, *supra* note 61, at 1380-81 (stating that the main justification for the creation of a property right in personal information is the utilitarian theory, which focuses on the maximization of benefit to all). When examining the problematic results of the information exchange (as discussed below), it is questionable whether the creation of such a right will lead to an optimal allocation of resources, thus defeating the cause for the creation of the property right. *See* CATE, *supra* note 60, at 74 (noting that if anyone, the collectors have a right in the information they compiled). On theories of property in general, see JOHN G. SPRANKLING, UNDERSTANDING PROPERTY LAW (2000).

66. For a description of various sources of personal information available in public records, as well as the problems this creates and possible solutions see Robert Gellman, *Public Records: Access, Privacy and Public Privacy*, *at* http://www.cdt.org/privacy/pubrecs/pubrec.html (last visited Aug. 20, 2003).

67. For example, demographic information regarding the residents of specific areas are combined with data regarding the specific individual's zip code. For more on this issue, see Latanya Sweeney, *Weaving Technology and Policy Together to Maintain Confidentiality*, 25 J.L., MED. & ETHICS 98, 100 (1997).

68. *See, e.g.*, *The Census and Privacy*, *available at* http://www.epic.org/privacy/census/ (last visited Sept. 25, 2003).

69. For example, see DM NEWS, *at* http://www.dmnews.com/ (last visited Nov. 21, 2002) (containing ads for the sale of lists created in the automobile, tool, and home brewery markets).

quirements could be enforced by the property regime mentioned, or by appropriate default rules.[70] To a certain extent, transactions on these secondary markets are already regulated by specific legislation,[71] yet in general, these transactions are permitted and widely practiced.[72]

Several companies, in their privacy policies, voluntarily address the regulation of secondary sales. The Federal Trade Commission (FTC) attempts to ensure that such companies enforce their own privacy policies as part of its role in ensuring fair business practices.[73] While these efforts on behalf of private entities and the FTC are important, they are far from sufficient. First, the FTC lacks the manpower to tackle such a feat and is unable to track down all those who breach their own privacy standards. Second, the FTC's reach with regard to such enforcement goes only as far as the company sets its standards.[74] Third, such an enforcement scheme will not allow the formation of an overall standard of conduct, but will depend on commercial entities to set the pace.[75] Finally, in many cases, the provisions presented in privacy policies tend to be quite hollow (at times including empty statements such as "We will share your information only with selected advertisers.").[76] After addressing the various aspects of the collection process, this Article returns to the description of the information flow and proceeds to describe the analysis stage.

B. Analysis

In this second stage, the data collectors analyze the information that was previously gathered in any one of the collection methods described above. It is at this stage that raw data turns into knowledge by way of meaningful patterns that provide collectors with insights into the present and future behavior of individuals. To properly understand both the analysis stage and analysis-based solutions, the Author must briefly return to examine the information flow paradigm and the location of analysis within this larger process. Because analysis constitutes the second phase of the information flow, its effectiveness relies on the quality and quantity of data in the previous collection stage. Thus, regulations adversely affecting the earlier stage of the information process (collection) will indirectly affect the

---

70. On these issues, see Kang, *supra* note 17 and Schwartz, *supra* note 17.

71. 15 U.S.C. § 1681(b) (2001). According to the Fair Credit Reporting Act, credit information could be passed on only in specific situations. *Id*. *See* Murphy, *supra* note 35, at 2410 (providing a list of other specific laws that regulate such secondary transactions including insurance, student information, financial data, and video tape rentals).

72. In the case of *Avrahami v. U.S. News & World Report, Inc.*, No. 96-203, 1996 WL 1065557 (Va. Cir. Ct. June 13, 1996), Mr. Avrahami sued, under a Virginia statute for the sale of his name to other databases, but was unsuccessful. For a description of these proceedings, see GARFINKEL, *supra* note 13, at 178-80 and Safier, *supra* note 9, at ¶ 102.

73. For a description of the actions taken by the FTC against GeoCities and Toysmart.com regarding their violation of their own policies, see Steven Hetcher, *The FTC as Internet Privacy Norm Entrepreneur*, 53 VAND. L. REV. 2041 (2000). *See also* Froomkin, *supra* note 10, at 1524 (discussing self-regulation).

74. *See* Steven Hetcher, *Changing the Social Meaning of Privacy in Cyberspace*, 15 HARV. J.L. & TECH. 149, 176-84 (2001). Hetcher discusses at length the market dynamics for privacy policies and why they do not result in a better situation for the collectees. *Id*.

75. *See id*.

76. *See id*. at 176-79.

analysis stage as well.

As a preliminary stage of analysis, the data collectors create "data warehouses" in which information from various sources is aggregated, sorted and organized to facilitate future access and analysis. There are several regulatory schemes focused on this intermediate process. The FTC's initiative, as well as other solutions, calls for an individual's right of access to these information warehouses which allow individuals to search for flaws in the data that pertains to them, as well as a right to correct outstanding errors.[77] Such requirements are also part of the rules governing the Credit Bureaus (major players in the personal information market).[78] Other regulatory schemes require the registration of databases of a certain size by a public registrar, as part of a larger scheme to oversee their management.[79] These regulatory steps are indeed important, yet fall short of providing solutions to the challenges society currently faces.

The next step in the data flow is the actual analysis of the organized personal information to obtain meaningful insights about individuals, both specifically and as parts of groups and subgroups. The analyst, using various software applications, probes the datasets in search of patterns and correlations, which will assist the relevant firm in future dealings with its present and future customers. This is done by constructing a hypothesis, which the analysts try to verify or falsify by testing the company's database. An example of such analysis is the much-discussed practice of "niche marketing," during which the analysts try to construct various classes of customers (the young-spending-single-professional, the retired-wealthy-widow, etc.), search for specific trends of behavior within every sub-group and thereafter divide the entire customer list according to these classes.[80]

Such methods have proved very useful in the past. Today, however, when collectors are faced with large and elaborate databases and have sufficient funds, they choose to utilize data mining techniques in order to obtain better results.[81] Such applications employ algorithms to reveal association rules and clusters within the data that might not have been apparent to the analyst initially sifting through the information.[82] As described elsewhere, data mining applications do not require a hypothesis to commence the analysis, but run through the database in search

77. "Access" is one of the fair information practices as well. Fair Information Practices, *supra* note 52.

78. Civil Liability for Willful Noncompliance (Fair Credit Reporting Act), 15 U.S.C. § 1681n (1998). *See also* ROBERT ELLIS SMITH, BEN FRANKLIN'S WEB SITE (2000) (providing an analysis of this Act).

79. For example, this is a requirement according to the EU Directive. FRED H. CATE, PRIVACY IN THE INFORMATION AGE 124 (Brookings Institution Press 1997). This is also a requirement according to Israeli law. See THE LAW OF PRIVACY PROTECTION 5781 § 8 (1981) (HTSHMA). In Great Britain, a similar requirement exists. Ronald J. Krotoszynski Jr., *Autonomy, Community, and Traditions of Liberty: The Contrast of British and American Privacy Law*, 1990 DUKE L.J. 1398, 1409 (1990).

80. *See* DAVID SHENK, DATA SMOG, 113-14 (1997).

81. *See generally MYOB*, *supra* note 1.

82. *See generally MYOB*, *supra* note 1.

83. *See generally MYOB*, *supra* note 1 (offering a technical description of the data mining process). For an extensive discussion of data mining applications see Usama M. Fayyad et al., *From Data Mining to Knowledge Discovery: An Overview, in* ADVANCES IN KNOWLEDGE DISCOVERY AND DATA MINING (Usama M. Fayyad, et al. 1996) and DAVID HAND ET AL., PRINCIPLES OF DATA MINING (2001).

of meaningful correlations between variables.[83] Therefore, the results of such analysis are unpredictable to both the collector and the collectee. Data mining allows collectors to reveal more patterns and correlations while using less manpower.

Of the three steps of the information flow, the analysis step is the only one that is not absolutely essential in the flow of information. When faced with small projects and databases, it is possible for an organization to leave out the in-depth analysis (and obviously the use of data mining applications) and implement the collected information directly. However, for most entities, given the immense amounts of data collected pertaining to many factors (information which now amounts to terabytes[84] of data per year), a casual glance at their databases will provide nothing but a headache. In today's competitive markets and especially in the Internet environment, the use of data analysis and perhaps even data mining is no longer a luxury, but an essential requirement.

When the collectees submit personal information, they usually do not intend to facilitate an in-depth analysis of their personal information, but to meet another objective. Therefore, during the analysis stage, the data is in many cases utilized for a different objective than the one for which it was originally submitted. In other words, the data is put to a secondary use. For example, when purchasers provide their credit card company with the details of a transaction, they intend the company to use this information to pay for the transaction and charge them for it at a later time. They probably do not intend to provide information that facilitates a complicated analysis of purchasing patterns and marketing strategies.[85]

The collector's ability to engage in secondary uses is central to the debate and regulation of the personal information process in general, and the analysis stage in particular. Practically speaking, limiting or prohibiting secondary uses will deliver a harsh blow to the analysis stage as a whole, since many of the analysis practices described will not take place. Yet such limitations indeed exist. For example, the European Union (EU) Directive prohibits any use of personal information beyond the specific use for which it was originally provided, without the consent of the relevant individual.[86] This requirement is echoed (to a certain extent) in some privacy policies adopted by commercial entities on their own accord.[87] In the United States, such secondary uses have been conditioned by the collectee's consent in several circumstances.[88] However, the regulation of secondary uses tends to be narrower than that of secondary sales (a controversial policy which promotes consolidation)[89] and in general is not accepted in the United

84. Tera = trillion, or 10^12. For example, Wal-Mart had 11 terabytes of transactional information for the year 1998. HAND, *supra* note 83, at 11).

85. *But cf. infra* Part II (discussing the possibility that the users are indeed interested that the credit card companies use their personal information to form patterns of their consumer behavior to allow them to detect fraud).

86. *See* EU Directive, *supra* note 4, ch. II, sec. 1, art. 6(1).

87. *See* Hetcher, *supra* note 74, at 181.

88. For example, such a right exists under COPPA, 15 U.S.C. §§ 6502(b)(A)(ii), (B)(iii) (2003). *See* Anita Allen, *Minor Distractions: Children, Privacy and E-Commerce,* 38 HOUS. L. REV. 751, 763 (2001).

89. For example, the Gramm-Leach-Bliley Act does not restrict the sharing of information among affiliates, but regulates the sharing of personal data outside the conglomerate. Edward J. Janger & Paul M. Schwartz, *The Gramm-Leach-Bliley Act, Information Privacy, and the Limits of Default Rules*, 86 MINN. L. REV. 1219, 1226-27 (2002).

90. SWIRE & LITAN, *supra* note 4, at 28 (stating that this provision is "the most surprising provision in American eyes").

States business arena.[90]

The regulation of secondary uses of personal data faces several difficulties. At first, there is an analytical-conceptual difficulty as to the rights of the collectees in the future uses of their personal information. In other words, it is not clear why collectees should control future uses of information after its collection in a non-intrusive manner. Shouldn't the collectors have the right to use such information as they please? Placing impediments over the right to use information may have free speech and First Amendment implications as well, since this is a restriction on the collectors' right to process and produce information (a derivative of the right of free speech). These difficulties are partially resolved by the property and contractual paradigms discussed above, as they provide the collectees with initial rights in their personal data, that are breached by the collectors' secondary uses.[91]

In addition, there is the issue of enforceability. Unlike the collection stage, which is, in many cases, carried out in the open, the analysis process is carried out internally, within the collectors' research departments. Therefore, it will be hard to track the extent of the actual analysis, and rules pertaining to such analysis will be quite difficult to enforce. Governmental enforcement might be substituted with providing individuals with the right to take private action against those carrying out such secondary uses. Yet again, such self-help provisions might be inappropriate for this situation.[92]

Lastly, obtaining the collectees' consent to such practices could still easily circumvent the secondary use-based regulation. The requirements for such consent, whether based on "opt-in" or "opt-out," or whether it must detail all the future uses of the information, are again a fertile ground for debate and disagreement. The use of data mining applications adds an additional complication to these debates. Since data mining analysis is not hypothetically driven and produces unpredictable results, it is nearly impossible to correctly present all the future uses of the personal information.

Beyond the regulation of secondary use, there are other possible regulatory schemes focused in the analysis stage. One such solution consists of prohibiting the use of specific variables as part of the analysis stage. Regulations mandating this solution may be put in place with regard to problematic factors such as the collectee's race, gender or sexual orientation. The rationale behind these rules is to block the collector's ability to discriminate between various individuals on the

91. However, this Author believes that the use of the property paradigm to provide for vested rights in the proceeds of secondary uses of personal information is insufficient. A prominent theory of property and ownership justification is the "labor-desert" theory. According to this theory, a property right is provided through the investment of labor. *See, generally* SPRANKLING, *supra* note 65. Applying this rationale to the issue at hand is challenging as it leads us to believe that the gathering of information should indeed create a property right—but in the hands of the collectors who went to many lengths to obtain this information. Even if the collectors will not receive a property right in view of their collection efforts due to their collection, certainly the analysts of personal information should receive a partial property right in the profiles and patterns they discovered. Clearly, this issue as well, required additional research.

92. This form of enforcement is ineffective because the public lacks information and understanding regarding these issues. In addition, such issues are not suited for class actions, since the damages are very hard to prove and are often too low. *See* Froomkin, *supra* note 10, at 1529; and Loomis, *supra* note 2 (discussing cookie-related litigation).

basis of these factors—conduct that is clearly contrary to public policy.[93]

The collection and analysis of personal information is only carried out to enable the final part of the process—the implementation and use of the personal information.

C. Implementation

The third and last step in the information process includes all actions benefiting from the collection and analysis stages described above. In today's business environment, the knowledge derived from analyzing personal information is used in many different settings. Companies use the feedback they receive from their customers to adjust their business plans, marketing initiatives and even research and development strategies. This paper focuses on the use of personal information that pertains to specific individuals, rather than general information pertaining to larger segments of society, thus addressing only part of today's many uses of customer-related information in the private sector.

In an earlier article, the Author describes a wide range of uses stemming from the analysis of personal information.[94] Possible uses include: (1) the targeting of customers with tailored advertisements and direct mailings based on their personal information; (2) the ability of marketers and vendors to discriminate among different individuals on the basis of their personal information and (3) the use or exposure of personal information with the intention to blackmail or embarrass the individual. In addition, recent trends of convergence in the media market lead to the danger of the "autonomy trap"—the ability of content providers to influence the opinions and conceptions of individuals by providing them with tailored content based on the provider's agenda and the individual's personal traits.[95]

The opportunities to carry out these tasks have expanded in view of the two recent phenomena discussed throughout this Article—the emergence of the Internet society and the availability of data mining applications. On the Web, the interface with specific clients could be tailored to their specific profile with simplicity and discretion as part of the "one on one" marketing scheme, which is an ideal setting for price discrimination.[96] Advertisers could easily provide customers with individually tailored banners and create additional opportunities to manipulate the public. Content providers will also benefit from this environment, as they now could provide individually tailored content (also referred to as the "Daily Me"),[97] and engage in personal agenda-setting initiatives (thus influencing the public even more than they are today). The growing use of data mining tools adds sophistication and precision to the implementation process, allowing the collectors to reach elaborate and surprising conclusions about their customers (even if such custom-

93. For a discussion of how such practices of discrimination take place in several situations, see *MYOB*, *supra* note 1 at *24. The EU directive also addresses this issue in Article 8(1). *See* Cate, *supra* note 4, at 433-34; *see also* SWIRE & LITAN, *supra* note 4, at 30.

94. *See MYOB*, *supra* note 1, at *17-48.

95. *See* Paul M. Schwartz, *Internet Privacy and the State*, 32 CONN. L. REV. 815, 821–828 (2000). Schwartz describes one of the traits of the "autonomy trap" as the "reduced sense of the possible." *Id*. at 825. In an earlier article, this term was used in a somewhat broader context to describe the ability of content providers to manipulate users by providing them with personally tailored content. *MYOB*, *supra* note 1, at n.106.

96. In addition, the Internet enables a leap (both in quality and quantity) in the ability to conduct individual surveillance and data collection.

97. This concept was introduced in NICHOLAS NEGROPONTE, BEING DIGITAL (1995).

ers may have tried hard to hide them) and use this knowledge to their benefit.

Here again, it is important to note the positioning of this stage within the larger process. Clearly, regulation of the earlier stages of collection and analysis will assist in solving the problems that result from the use and implementation of personal information. However, we must not ignore the possible downside. Strong regulation of these earlier stages will leave fewer opportunities for the use of personal information, and all the benefits that it might provide.

The implementation stage is surprisingly neglected when solutions and proposed regulations are debated. A review of the few solutions that are rooted in this final stage of the data flow may prove useful.

In the context of mass marketing and telemarketing, implementation solutions can be viewed as the formation of "do not call" lists, allowing customers to obtain an unlisted number and permitting customers to "opt out" of the mass mailers' mailing lists.[98] In the context of "spam" mail (which differs from other forms of mass mailings in that the marginal costs to the marketer are almost zero) there have been several proposals as well as actual state legislation restricting specific actions of mass e-mailers.[99] In addition, technological solutions have been developed to filter out spam and junk mail.[100]

Several of the classic privacy torts are aimed at regulating the implementation stage. The torts of Publicity or Seclusion aim to counter the injury from harmful use of personal information.[101] The tort of "misappropriation" protects individuals from commercial uses of their name or likeness—again regulating the possible uses of personal information.[102]

In the Internet context, a solution recently suggested consists of providing customers using a "market for one" application with a glimpse of the information the vendor has collected about them and how it is currently used. This application, "the cookie viewer" (discussed at length below), will show the customers what segment or cluster the vendor has placed them into.[103] Thus, it will provide some insight into the vendor's practices and shield the users from potential manipulation.

In addition to the solutions mentioned above, there are other options to be explored within this stage, such as new regulations aimed at curing the specific harms stemming from the practices of price discrimination and the autonomy trap.

---

98. *See* Sovern, *supra* note 42, at 1313. According to Sovern, such rules have been introduced in several states. *Id.* at 1315-17. At the Federal level, *see also* The Telephone Consumer Protection Act of 1991, 47 U.S.C. § 227 (2001); Telemarketing and Consumer Fraud and Abuse Prevention Act, 15 U.S.C. §§ 6101-08 (1998 & Supp. 2003); and the FTC's Telemarketing Sales Rules, 16 C.F.R. pt. 310 (2003).

99. *See* Rob Gavin, *E-Commerce (A Special Report): The Rules,* WALL ST. J., Oct. 21, 2002, at R9 (discussing recent Spam-related legislation in various states including California).

100. *See* James Gleick, *Tangled up in Spam*, THE N.Y. TIMES MAG., Feb. 9, 2003, at 42.

101. *See* RESTATEMENT (SECOND) OF TORTS § 652D (1965). *See also* Solove, *supra* note 4, at 1433 (analyzing this tort and leading to the conclusion that plaintiffs rarely win) and Zimmerman, *supra* note 21.

102. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS, § 652C (1977). *See* discussion *supra* note 40.

103. Such solutions have been suggested as part of the recent FTC-DoubleClick settlement. *See* discussion *infra* Part III (discussing the "cookie viewer"). *See generally* Tal Zarsky, *Cookie Viewers and the Undermining of Data-Mining: A Critical Review of the DoubleClick Settlement*, 2002 STAN. TECH. L. REV. 1 (2002).

By enacting effective regulation at this stage, the problematic consequences of the earlier stages could be minimized.

D. Summing Up

The flow of information does not stop at the end of the process described above, but continues in an everlasting cycle. After the collectors use the personal information in their interaction with their customers, they continue to collect their responses. Such responses are later analyzed and used again when interacting with the same and other customers. This feedback cycle is very helpful to collectors, since it enables them to assess the accuracy of their original analysis and implementation practices and allows them to adjust their actions accordingly, until reaching a satisfactory level of precision.

When taking an overall view of the solutions spread along the information flow, it is apparent that most of today's debate is concentrated in the collection stage, with less emphasis on the final implementation stage. At first blush, this is a surprising result—shouldn't solutions focus on the stage in which the problems actually occur and where the damage is actually caused? This outcome results from the widespread belief, held by many of today's legal scholars, that the collection of information itself is a problematic practice. This well-accepted, yet somewhat purist, perspective asserts that mere collection violates the rights of individuals and should be stopped, regardless of the future uses of such data.[104] Scholars state that surveillance on its own merit could adversely affect an individual's state of mind and inhibit daily activities.[105] Other leading legal thinkers assert that knowledge of constant surveillance will promote conformity and intrude on the individual's autonomy.[106] Therefore, regulating the latter stages of the data flow is insufficient.

Yet even though these claims may seem convincing, the Author begs to differ. The fears stated as part of this "purist" position indeed seem troubling, yet society will have to adapt (and is indeed constantly adapting) to the changes in technology. These fears are too opaque, abstract or far-fetched to require actual changes in the information flow.[107] Since any form of regulation within this context will be met with strong and powerful opposition, legislation must be pragmatic and focus its concern on the actual detriments that may occur. Solutions should protect the public from dangerous uses of personal information, rather than mere surveillance.

In the following sections, this Article argues that contrary to widespread belief, regulation of the personal information flow should focus on the implementa-

---

104. Interestingly, the origins of the Jewish law's concepts of privacy are based on this concept. *See* JEFFREY ROSEN, THE UNWANTED GAZE: THE DESTRUCTION OF PRIVACY IN AMERICA 19 (2000).

105. Solove, *supra* note 22, at 1130. *See also* Cohen, *supra* note 61, at 1425 (stating the importance of having an unmonitored choice); Kang, *supra* note 17, at 1212-17 (stating that monitoring on its own should be restricted, in order to avoid embarrassment, enable the construction of intimacy, and avert misuse).

106. Solove, *supra* note 22, at 1130 n.247.

107. In addition, such claims are somewhat tainted by today's media reality that presents many people who go to great lengths in order to expose themselves to the public eye, while seeking publicity.

tion stage. This conclusion is premised on several critiques of the collection-based solutions, which are detailed below. The problems described above and elsewhere as stemming from the collection and use of personal information could be answered directly, rather than impede on the entire collection process. These arguments are further strengthened when one considers the widespread use of data mining tools, forcing us to look at the previously presented solutions with a new perspective.

II. PREFERRING THE REGULATION OF IMPLEMENTATION TO COLLECTION

A. Virtual Babies and Virtual Bathwater

When focusing regulation on the earlier stages of collection and analysis, obvious side effects are created by severely reducing the quality and extent of the possible analysis of personal information. This will surely occur in view of changes in tort liability, the creation of property rights in personal information and the change of default rules to those that provide for limitations on information sale and analysis. This result is far from optimal. Society has much to gain from the collection and analysis of information (perhaps the most significant asset of our time) and a great deal can be derived from its availability and use.[108] There are several concrete examples of the benefits derived from the collection of personal information, including a vibrant information market and data analysis. These examples prove that collection and analysis of personal information (as practiced today) should be permitted, thus urging us to concentrate the regulatory effort within the last stage of the information flow. These especially point out the benefits of secondary uses and secondary sales, which are frequently addressed as issues that must be regulated and limited.

1. Subsidies for Startups and the Importance of Innovation

The ability to collect and use personal information is, without a doubt, beneficial to all collectors, who use such information to enhance their marketing initiatives.[109] However, access to such information is especially helpful to new companies.[110] Not only are new companies struggling with the expenses of research, development and starting their business, they must also break into the public's awareness. By using personal information, a new company can close the marketing gap between it and large incumbents. Access to personal information enables startups to directly contact specific individuals who might be interested in their product, rather than engaging in a wide marketing campaign at a much higher cost. Through analyzing the collected personal information, new companies can obtain indications as to the specific type or profile of their ideal customer and can contact such individuals directly. Using such analysis, new companies could launch a

108. On the importance of balancing privacy rights in view of other benefits and interests, see generally AMITAI ETZIONI, THE LIMITS OF PRIVACY (1999).

109. SWIRE & LITAN, *supra* note 4, at 77-79 (stating that a strong privacy rule will inhibit innovation). *But see* NEIL POSTMAN, TECHNOPOLY—THE SURRENDER OF CULTURE TO TECHNOLOGY (1993) (viewing constant technological progress as not being a plausible and positive outcome).

110. *See* CATE, *supra* note 79 at 14; *see also* FTC report, *supra* note 52 at 32.

cheap and minimal marketing initiative with a very high chance of success. Data mining applications, already widely in use, can greatly enhance these analysis capabilities and marketing initiatives. Data mining provides for more accurate predictions, as well as an improved ability to cluster groups of customers with similar interests together. This all enables the startup companies to run less expensive and more effective marketing campaigns.[111]

Providing a "startup friendly" environment is in the best interest of us all. An environment that accommodates the introduction of new ideas and products facilitates real competition in which the best product emerges a winner, regardless of whether it is promoted by an incumbent or a new company. A startup friendly environment will also quicken development of newer and better technologies and generally promote innovation. Creating a business environment that is friendly to new companies is of even greater importance given the traits of the Internet and the telecommunications business landscape. In these settings, early movers gain a great advantage; many of the first websites to offer services are still the leaders of the online industry (Amazon.com, Yahoo!, Google, Ebay and others).[112] Such advantages probably result from the Internet's infrastructure and the way Internet users develop their surfing habits.[113] In addition, network effects are very common in certain parts of this environment, and breaking into the market (even with a superior product) can prove extremely difficult.[114] Therefore, allowing information collection and analysis will encourage innovation by smaller companies in a somewhat hostile business environment. In view of this analysis, it is clear that many of the proposed regulations and limitations of the collection process will benefit large incumbents (who will in turn be very interested in such rules). Therefore, we must be cautious of the hidden agendas of those trying to promote privacy by inhibiting collection, as they might be pursuing this objective to protect the dominance of incumbents against smaller, startup companies.[115] Preference should be given to regulation that allows for a vibrant information market, while at the same time protecting the interests of the customers rather than impeding collection from the outset.

An example of a regulatory scheme put in place to assist new companies in a hostile environment is the United States Telecommunication Act.[116] The regula-

111. *See* PETER CABENA ET AL., DISCOVERING DATA MINING—FROM CONCEPT TO IMPLEMENTATION 123 (1998) (describing a company that through the use of data mining tools could send out fifty percent fewer advertisements, and still reach ninety percent of its best prospects).

112. Several companies created dominant brands in the Internet environment within a relatively short time. Paul Judge, et al., *The Name's the Thing*, BUS. WK., Nov. 15, 1999 at 36.

113. This might be a result of "bookmarking" or adding sites used to the "favorites" lists, which are later solely relied upon. In addition, on the Web, the user cannot "look out the window" and notice new stores and establishments, but goes to specific addresses and destinations.

114. For example, note the difficulty of Linux and Apple versus the Windows operation systems. In addition, note the current battle in the "instant messaging" arena for setting a protocol. Generally, "network effects" refer to situations in which the consumer's decision will depend on the expected growth of a network. Competition in a market with network effects is problematic, as the quality of the network depends directly on the number of subscribers. Therefore, consumers are reluctant to join new networks. For a discussion of these issues, see Robert D. Anderson & Nancy T. Gallini ed., COMPETITION POLICY AND INTELLECTUAL PROPERTY RIGHTS IN THE KNOWLEDGE-BASED ECONOMY, 228 (1998).

115. FRED H. CATE, PRIVACY IN PERSPECTIVES 14 (2001).

116. Telecommunications Act of 1996, 47 U.S.C. § 251(c) (2001).

tion of the telecommunications market introduced "must carry" rules that force incumbents (ILEC's) to provide startups (CLEC's) with access to their physical infrastructure.[117] Such regulation was required since the initial expenses for new companies (especially in creating the "last mile" connection to the homes of the customers) were too high for a healthy competitive market to emerge.[118] In a broad analogy to the personal information market, the "last mile" in this context can be viewed as the virtual last mile of specific personal information that connects sellers to potential buyers. This last mile can be paved by using personal information to connect the potential clients with the new startup companies.[119] With access to this virtual last mile, new companies will have an easier start in their uphill battle.

The importance of creating an environment that increases innovation has recently been the subject of public debate, yet from a different angle—that of intellectual property. In *The Future of Ideas*, Professor Lessig argues that recent changes in the laws governing intellectual property (with a focus on copyright) will decrease the amount of such property in the public domain, or available for personal uses through other exceptions.[120] This, in Lessig's opinion, will cause a serious inhibition on innovation.[121] Since innovation is viewed as an important force that generates growth and progress, Lessig strongly argues against these changes in the law.[122] The analogy to the issues at hand is clear; innovation is an important aspect to keep in mind when contemplating which public policy to adopt in the context of the personal information market as well. Allowing the collection and analysis of personal information will enhance innovation and should be allowed to the greatest extent possible.

2. Creating Value

Beyond the specific benefits to small businesses, permitting collection and analysis of personal information by focusing regulatory efforts on the implementation stage allows the entire public to directly benefit from the analysis of personal information. Analyzing personal data provides new information and insight that can be utilized to the benefit of all, and therefore should not be lost. The ability to analyze this information using data mining applications opens a portal to

117. In *Verizon v. Iowa Utilities Bd. et al.*, 535 U.S. 467, 475 (2002), the court states that in accordance with the Telecom Act: ". . . new entrants are entailed among other things, to lease elements of the local telephone network from the incumbent monopolist." In addition, the court presents the reason for this regulation: "At the dawn of modern utility regulation, in order to offset monopoly power and ensure affordable, stable public access to a utility's goods or services, legislatures enacted rate schedules to fix the prices a utility could charge." *Id.* at 477.

118. *Id.* at 490. The court states that "[a] newcomer could not compete with the incumbent carrier to provide local service without coming close to replicating the incumbent's entire existing network, the most costly and difficult part of which would be laying down the 'last mile' of feeder wire. . . ." *Id.*

119. The Author concedes that the situation described in the telecom market is very different, as in this setting the ILEC's are monopolists that have in many cases received the original licenses and property for free or at a very low price, and thus the government has proper justification to force them to "carry" information sent by others. However, the requirements described in the information market are not nearly as harsh and do not negate any existing property right.

120. LAWRENCE LESSIG, THE FUTURE OF IDEAS, 234-39 (2001).

121. *Id.*

122. *Id.*

additional sophistication, deeper patterns of extended knowledge and adds value to society. Generally, the availability of personal information for analysis and data mining can promote various forms of scientific discovery. This is best demonstrated by the various Global Information System (GIS) projects that form detailed maps, which include ecological, geological and other data together with the personal data of the residents of a particular area.[123] The analysis of a GIS database can lead scientists to important discoveries that improve health and quality of life. For example, such analysis finds correlations between diseases and areas of residence and thus can reveal the causes of these illnesses.[124]

However, most criticism of the process of collection and analysis of personal information is not directed at research carried out by the scientific community, but at uses within the business arena. Yet here too, such data analysis proves to be important. Fraud detection and the cutting of credit costs are uses of personal information that emphasize the strengths of such analysis.

Analyzing consumer behavior, especially in the context of credit card use, allows collectors to identify purchasing patterns for an individual. A transaction that deviates from these patterns is flagged as possible foul play. The use of these methods assists law enforcement in capturing criminals (especially those involved in identity theft) and assists in cutting costs for the cardholders and insurers by reducing the expenses associated with such crimes.[125]

Cutting Credit Costs is another issue often mentioned as a benefit of personal information analysis. This task is carried out by analyzing personal information (at times, by using data mining applications) to assess credit worthiness.[126] Broad access to information and enhanced analysis abilities allow for better assessment of the underlying risk in providing loans and extending credit. As a result of such analysis, lenders improve their ability to assess the risk associated with every loan as well as their pricing schemes for financing and to avoid clients that default on payments. The availability of such information and technology led to a decline in the overall costs of credit (as defaults were eventually covered by other borrowers).

Even though these examples seem convincing and are popular as counterclaims to those who advocate strong privacy regulation, something is still missing. Fraud detection and credit rating are not the mainstream uses for personal information but are the extreme examples, where the benefits are clear and people tend to voluntarily submit their personal information to receive such services. These examples do not sufficiently reveal the tension between the benefits and dangers that could arise from the analysis of personal information.

The retail and content markets present the best opportunities to create value for both seller and buyer through the data mining of personal information. Here, data mining applications can reveal hidden patterns and correlations in consumer

123. "GIS" stands for Geographic Information Systems. It is defined as "a system of computer software, hardware, and data, and personnel to help manipulate, analyze and present information that is tied to a spatial location." GIS.COM, *Geographic Information Systems "GIS": Geography Matters*, at http://www.gis.com/whatisgis/whatisgis.pdf (last visited Sept. 13, 2003).

124. Such research is to a great extent made possible by mandatory tumor registration carried out by hospitals. Ron Breyer, *Limits to Privacy*, Columbia Bioethics Center (Feb. 20, 2003).

125. CATE, *supra* note 115 at 11.

126. CATE, *supra* note 115, at 11.

127. "Hidden" does not necessarily mean it was actively hidden by the customers, but that it is not apparent to the vendors.

behavior, which can later be used to the benefit of both parties.[127] The most famous example of the benefits of the analysis of personal information in this context is the correlation between the purchase of beer and diapers on a late Friday afternoon.[128] Even though this correlation is only a hoax, it has sparked the imagination of many marketers and businessmen who strive to reveal patterns within consumer behavior, which have gone unnoticed by customers and vendors. Using such knowledge, sellers can rearrange their shelves and change their marketing schemes so as to maximize their revenues. When consumers find the product they are seeking more quickly, the vendor benefits from the sale and creates goodwill, while the buyers save time and effort. Yet even the "beer and diapers" paradigm refers to analyses that utilize only general information (regarding the number of products bought at specific times, or a simple "basket analysis") and therefore fail to encapsulate the benefits of personal information analysis. To find the business practices where actual personal information about individuals (as opposed to personal information about groups or sub groups of individuals) is utilized, we must take a closer look at advertising and direct marketing.

Today's consumer markets are growing, as are the variety of products and the mediums to advertise such products. These expansions force vendors to spend substantial funds on excessive advertising campaigns. Obviously, the consumers pay a portion of these expenses as part of the product's price. However, by analyzing personal information, vendors can make strong predictions as to whom might be interested in their product. Vendors therefore can focus their marketing initiatives on this limited crowd and avoid spending funds on pointless campaigns.[129] As a result, vendors can cut down on advertising and operating expenses, eventually lowering the price the customer is charged for the product.[130]

Beyond the simple advertising perspective, marketers can use customer lists and data mining to create value in various ways. For instance, they can inform their customers which of their favorite items are on sale in a given week, or even contact them directly when they are stuck with a surplus in their inventory. That way, the consumers can purchase the product they need at a low price, while the vendor will not be required to spend resources on storage or even lose the products (especially with regard to perishables). Such results would not be possible without access to the customer's personal information and therefore, preferences, and the ability to reach out and contact the individual client.

An example of applications that can create value for both consumers and vendors are recommendation systems. These tools, now offered by several web sites and usually powered by data mining algorithms, provide the website patrons with recommendations for products that are most likely to interest them given their specific profile.[131] These systems make suggestions based on the partial informa-

128. *See* http://web.onetel.com/~hibou/Beer%20and%20nappies.html (last visited Sept. 19, 2003). A new example that is based on case studies carried out by Wal-Mart mentions the increased sales of Band-Aids after placing them in close proximity to fishhooks. Matt Richtel, *Applying Science to the Casino*, N.Y. TIMES, Nov. 3, 2002 at 2.

129. FED. TRADE COMM'N, *supra* note 52.

130. For a thorough analysis of these points, see Kent Walker, *supra* note 53 at 89-92. Walker's analysis mentions the value of information collection, both with regard to the single consumer, and on the macro level with regard to the entire market.

131. *See, e.g.*, Amazon.com and www.Cdnow.com. For a discussion on the mechanics of such recommendation systems, see MYOB *supra* note 1, at *7, 9.

tion a customer provides while using previously constructed patterns, thus saving the customers' time and money. Recommendations, of course, benefit vendors as well, since a good recommendation increases the chances of a final purchase. Recommendation systems will continue to advance in sophistication so long as the databases upon which they draw keep growing. These systems will also retain the ability to "cross-recommend" products from different fields of retail or interest (for example, from movies to books to cuisine) on the basis of the elaborate patterns they construct.

In view of the positive prospects of innovation and creation of value, the use of personal information in conjunction with data mining can work to the benefit of both collectors and collectees, and everyone ends up smiling. If this is the case, why are the practices of personal information collection and analysis creating a general uproar in public opinion and by legal scholars? It is probably because of the dangerous flipside that might occur. The downside to the analysis of personal information is that rather than using its results to the benefit of all, the collectors will use this knowledge to their exclusive benefit, while manipulating and taking advantage of their customers. Instead of understanding and meeting a customer's demand, the information will be used to overcharge the customer. Instead of using this data to provide concise and pertinent content, content providers will choose to bombard our senses. Human nature will cause the beneficial practices to remain utopian, leaving us with a grim reality.

Yet even though such concerns are justified, they should be dealt with directly by regulating the implementation stage rather than placing restraints on the earlier stages of collection and analysis. By taking the implementation route of regulation, society can benefit from the advantages data mining provides and remain safe from manipulation. Let us not throw out the baby with the bathwater; by regulating in the early stages of the data flow, we may indeed solve the problem, but also eliminate many of the benefits of the technology. Note, however, that the arguments above do not intend to advocate a transparent society, or the full publication of personal information.[132] The assertion is merely that in order to protect society from the harmful outcomes of the use of personal information, the benefits of personal information analysis need not be lost. The emphasis is on the problematic outcomes of extensive restrictions on secondary markets and secondary uses, which are the basic tools for facilitating the initiatives described above. By maintaining the collection and analysis status quo and placing emphasis on implementation-based solutions, we can have the proverbial information privacy cake, and eat it, too.

B. The "Personal-Information-Based" Transaction

As described above, a popular form of solutions in general and of collection-

132. The benefits of publicizing personal information have been explored in detail from the "law and economics" perspective. This approach claims that it is economically efficient for personal information to be publicized so as to avoid searching costs and errors. This perspective clashes with humanitarian perspectives, which claim that the right to privacy cannot be reduced to an economic analysis. For a discussion of these issues, see Murphy, *supra* note 35 at 2381-82. *See also* Eli M. Noam, *Privacy in Telecommunications—Markets, Rights and Regulations*, NEW TELECOM Q. 4Q 1995, at http://www.tfi.com/pubs/ntq/articles/view/95Q2_A8.html (last visited Sept. 4, 2003).

based regulation in particular is the regulation of personal information-based transactions (due in part to the incompatibility of privacy and tort paradigms with the issues at hand). At first glance, such solutions seem promising. They do not extensively interfere with the flow of information, but introduce a market mechanism to facilitate a fair transfer of personal data. They also seem to provide a compromise that the opponents of strong privacy regulation might accept. Nevertheless, such solutions are flawed.

To make these transactions possible, the collectees are provided with property rights in the personal information that pertains to them.[133] Yet the rationale for creating these rights is complex. As with other property rights that are attached to intangibles (such as intellectual property), these rights are non-rivalrous, as they could be held by an unlimited number of parties at the same time.[134] Yet unlike other forms of intellectual property, such as copyright or patents, there is no reason to provide incentives for creating personal information since the creation of such information does not require any form of innovation, but simply encourages people to live their lives as usual.[135] There is also no reason to create a property right to promote the alienability of such information; the reasons for providing such a right are quite to the contrary. The dominant underlying reason for creating such property rights is to ensure that the collectees are sufficiently compensated for the damage they might incur as a result of the collection and use of their personal information. Another approach will suggest that these property rights are created to limit situations in which collectors gain from the collectees' information without compensating the collectees. Note that this second approach is somewhat circular, because without an initial right the collectees have no claim in the future profits of the collectors.

A market for this new commodity will surely follow the creation of an alienable right in personal information.[136] In this market, the collectors will provide collectees with various conveniences and compensation in exchange for their personal information. The price set at this market will be the equilibrium between the demand and supply price curves of the collectors and the collectees. The collectors' demand curve will reflect the benefits the collectors could derive from this commodity in its future analysis and use.[137] The collectees' supply curve will reflect the perceived detriments they might face as a result of providing this infor-

133. For a discussion regarding property rights, see *infra* Part I.

134. See LESSIG, *supra* note 120, at 131 for an analysis of this concept in the intellectual property context.

135. Cohen, *supra* note 61, at 1387-88.

136. *See* Lemly *supra* note 62.

137. The analysis of property rights fails to address the Hegelian or "personality" approach to property rights. Margaret J. Radin, *Property and Personhood*, 34 STAN. L. REV. 957, 957 (1982) (asserting that "to achieve proper self-development—to be a person—an individual needs some control over resources in the external environment"). Such control is asserted by the creation and use of property rights. *Id.* In light of this perspective, it is clear why property rights should be provided in personal information, as the personal information is directly connected to the owner. However, this Article omits this theory from its general discussion, as it does not add clarity or insight to the situations in which this new right is both alienable and subject to a transaction between collectors and collectees.

138. In addition, this price will be affected by the value individuals attach to surveillance in general. As mentioned above, people might object to mere surveillance, as it might inhibit individual thought, promote conformity, etc. *See* sources cited *supra* note 105.

mation and allowing its use.[138]

Yet, as shown in the following paragraphs, the outcomes of this market dynamic will prove to be undesirable. It will not fairly resolve the problems of information privacy and moreover, can lead to further inequities. The primary reason for such shortcomings is that these transactions take place at the beginning of the data flow process. The introduction of data mining applications to this field aggravates the situation and exacerbates the problems for reasons this Article will detail. These conclusions provide us with additional reasons to focus regulatory attempts at the latter stages of the information flow.

1. Myopia and Corrective Lenses

The outcome of the transactional dynamic described above will always favor the information collectors. The legal scholars suggesting these market solutions were obviously aware of this problem.[139] They therefore provide several suggestions to transform these transactions into "collectee friendly" ones by mandating notice and choice, changing the default rules and creating property rights. However, as with almost every transaction, the final word as to the specifics of the information based agreements will be that of the parties' themselves. Even though the collectees are provided with notice, choice, property rights and perhaps favorable default rules, they can still surrender their information privacy if they believe they are being properly compensated.[140]

Therefore, the success of regulatory attempts focused on the personal information transaction depends on whether the collectee can correctly balance the benefits and detriments of such a transaction, especially at the early stage of information collection.[141] The collectees must carry out such balancing tasks in order to establish their "supply" price that will eventually affect the final equilibrium. However, there is a valid argument that a fair transaction cannot take place at this early juncture, and that the results of such a transaction will always favor the collectors. This assertion is backed by several explanations. One such explanation points to the unequal information and understanding collectors and collectees will have regarding the uses and harms possibly associated with this transaction.[142] In addition, the collectees are faced with unbearably high transaction costs[143] when trying to reveal the future outcomes of submitting personal information. Several legal scholars even refer to the environment in which these transactions take place as one that consists of a market failure (since the collectees are unable to reach a

139. Froomkin, *supra* note 10, at 1533.

140. Froomkin, *supra* note 10 at 1535 (noting that in the EU, even though strict privacy rules have been implemented, individuals are still giving away their privacy rights quite easily).

141. The allocation of a property right in this context is problematic due to the lack of proportion in loss between collectors and collectees. A. Michael Froomkin, *Flood Control on the Information Ocean: Living with Anonymity, Digital Cash, and Distributed Databases*, 15 J.L. & COM. 395 (1996).

142. See Noam, *supra* note 132, for an economic analysis of transaction costs, symmetry of information and market failure.

143. Janger & Schwartz, *supra* note 89, at 1240 (addressing problems with similar information-based transactions through the "Lemon Equilibrium," where customers lack full information regarding non-price terms and are therefore unable to obtain fair results in negotiations).

144. Schwartz, *supra* note 17, at 1683, n.440 (providing an analysis of the reasons for such market failure).

fair agreement with the collectors, given the difficulties they are facing).[144] Similarly, Michael Froomkin refers to an inherent myopia[145] from the collectees' perspective, who are unable to grasp that the "whole" (the knowledge and benefits the collectors can derive from the information as well as the detriments the collectees might eventually suffer) is greater than the sum of all the "parts" (the various bits of personal information provided).[146] Collectees cannot correctly assess the value of their personal information at the time of collection, as the dangers are far away and indefinite at that time. Therefore, regardless of the various regulatory attempts, information-based transactions will favor the collector.[147] They will reflect a price that is too low both to cover the benefits to the collector or the detriments to the collectee. Solutions that unjustly enrich the collectors are not achieving their objective, and force us to rethink their overall effectiveness.

The collectee's myopia is exacerbated by the growing sophistication of data mining tools. With these tools, collectors can reveal additional hidden patterns, thus enriching their knowledge and increasing the possible uses of the information. Thus, the larger picture, featuring all the possible uses of the personal information after forming patterns and clusters, is driven farther and farther away, clearly out of the myopic collectees' range of sight. Given these abilities of enhanced analysis, the chances are quite slim that negotiations occurring in the collection phase will result in a price that is fair to the collectees.

To understand the complexity data mining adds to the information-based transaction, consider the following example:

When Al goes shopping at the local supermarket, he always makes sure to bring his "Club Card" with him so that he is able to save as much as one or two dollars on every purchase. At the same time, the supermarket gets something from Al—a great deal of personal information.

(a) From Al's purchases of low sugar products, the supermarket might deduce that Al is trying to go on a diet. Accordingly, the supermarket mails him advertisements for new diet pills.

(b) From the frequent yet limited purchases and the use of the delivery services, the supermarket deduced that Al does not own a car. Therefore, the supermarket knows that Al is probably unable to drive to a distant location for a better bargain and does not supply him with monthly coupons.

(c) From the purchases of baby food and diapers, the supermarket concludes that Al's family has a baby. Therefore, the supermarket informs a "Young Mother" magazine to send Al's family an offer to subscribe.

(d) From the purchases of dog food, the existence of a dog is assumed. In response, the supermarket sells Al's name to a dog food supplier.

As illustrated by these examples, collectors can use personal information in various ways. It is fair to assume that the collectees can indeed predict and anticipate some of these uses, and after balancing them against the reduced price they were offered, opt for the latter. However, can we expect Al to consider the fact that

145. Froomkin, *supra* note 10, at 1502.

146. Richard T. Ford, *Save the Robots: Cyber Profiling and Your So-Called Life*, 52 STAN. L. REV. 1573, 1573 (stating that both parties cannot anticipate the outcome of such collection).

147. For example, Forrester Research has concluded that the value individuals attach to their total privacy is about five dollars a month. *See* Bob Tedeschi, *Tech Briefing Internet: The Price of Online Privacy*, N.Y. TIMES, June 19, 2002, at C4. *See also* Saul Hansell, *The Big Yahoo Privacy Storm that Wasn't*, N.Y. TIMES, May 13, 2002, at C4. Yahoo! changed their privacy policy to the detriment of customers, yet very few reacted. *Id.*

the information collected, when aggregated with information gathered from other sources and previously constructed patterns, reveals that there is a 58% chance that he will consider having another child within eighteen months, a 64% chance that he will consider the purchase of a car in twelve months and 73% chance his wife will vote "Republican" in the upcoming elections? Probably not. Yet such predictions might be possible through the use of data mining applications and careful analysis of personal information. Data mining adds an additional dimension to personal information analysis, a dimension that is difficult to grasp at the time of collection.

In conclusion, at the time of the information-based transaction, the collectees are unequipped to bargain adequately for proper compensation for their personal information. Therefore, collection-based solutions premised on a transactional perspective should be avoided categorically, as in almost every case the collectees will receive the lower end of the deal.

There are two options to overcome this inherent difficultly. One option is replacing the somewhat flexible default rules and alienable property rights with stringent regulation. However, this option is not practical, given the information collectors' influence on legislators, nor advisable given the benefits that can be derived from the analysis of personal information.[148] It also does not conform to the general freedom of contract.

Alternatively, a solution could focus on solving this severe case of myopia by regulating the implementation stage. To carry through the optical metaphor, we could resolve the myopia problem by reducing the distance between the objects (the point of regulation and the actual use of the information), rather than applying corrective lenses.[149] Solutions premised on the information-based transaction do not only result in unfairness in the relations between collectors and collectees, but also among the collectees themselves.

2. Collectees vs. Subjects of Manipulation

The information-based transactions provide some compensation for the detriments that result from the collector's actions. As part of a market dynamic, such compensation will be proportional to the amount of information the collectee is conceding—the more information provided, the more compensation received. Yet is such an outcome desirable and fair?

To answer this question, let us revisit the personal information market's pricing dynamic. In this market, the price reflects various considerations of both the seller and the buyer. To the buyer, or collector, the price of personal information reflects the sharing of the benefits they are reaping from the personal information of others. On the other hand, the sellers, or collectees, are not only interested in their share of the benefits (which I refer to as "perspective (1)"), but are seeking compensation for the repercussions they might suffer as a result of such collection ("perspective (2)"). These perspectives seem to be aligned but are actually quite distinct. In the past, before collectors had retained their current ability to compute

148. Another argument against obligatory rules of inalienability is presented by Kang, *supra* note 17, at 1266 (stating that such rules will risk surrendering control over the right of privacy to the state).

149. *See* discussion *infra* Part IV.

and analyze data, distinguishing between these perspectives was trivial. The more information the collectee provided, the greater the benefit collectors could derive from it, and consequently, the more harm the same collectee might suffer. Data mining, however, adds a twist and makes these matters more complicated.

In order to understand the distinction between these two perspectives, consider the following examples:

(1) *George must get his beloved wife an anniversary gift. George does not use the Internet often, but nevertheless decides to visit shoeperstore.com (a fictional online woman's shoe retailer) to purchase shoes. Even though George is aware of the risks associated with online shopping, he agrees to log on and provide some personal information, since the site promised its shoppers a ten dollar rebate at the local bookstore. After George provides the e-commerce site with his personal information, the site installs a cookie in his hard drive and starts tracking his actions through the website and beyond. The site records that George spent very little time at every web page, did not click on any advertisement or notices regarding specials, sales or reduced prices, and purchased several low quality items at a high price. After analyzing this data aggregated with other consumer information, the website placed George in a cluster designated as the "rich, busy and does not understand" type of consumers.*

The next year, around the same date, George returns to shoeperstore.com website (which is no longer offering the bookstore promotion). The website identifies George and his previously determined shopping pattern and presents him with exclusively "high price-low quality" merchandise. George does not suspect this manipulation and again purchases overpriced items.

(2) *In this hypothetical, George visits the website only on the first year and does not return to shoeperstore.com, or any other online retailer. However, elsewhere in George's hometown, Ken is logging onto the same website for the first time. The website recognizes Ken (by means of cookie, etc.) and from information it obtained on the secondary market is able to learn that Ken too is a male, well-to-do and shares George's zip code. With such partial information, the website refers to a predictive model it created using information previously acquired from George and other shoppers. Based on this model, the website places Ken in the same category in which it previously placed George ("rich, busy and does not understand") and thereafter presents him only with overpriced products of poor quality. Ken, who does not suspect a thing, falls for the trap and purchases these overpriced items.*

Example (1) is simple: From the first perspective, that compensation is, in fact, part of sharing the benefits the collector derives from the data, shoeperstore.com is paying George for the use of his personal information by providing him with a ten dollar coupon. From the second perspective, that the compensation is for the future damages that such collection might cause, this transaction appears very much the same. George receives ten dollars and in exchange accepts the risk of any adverse outcome that may result from recording his personal shopping patterns, linking these patterns to his personal information and using such knowledge in future interactions.

The second example is quite different. Here, George provides the website with information that was harmful to Ken, due to the similarities in their personal attributes and behavioral patterns. When George was "telling" shoeperstore.com

about himself, he was in fact "telling" it about Ken as well. Therefore this second example allows us to distinguish between the two perspectives mentioned above. While the ten dollar coupon could be considered as the website's payment to George for future benefits derived from his personal information, it does not seem to properly compensate for any detriment the information might bring about. In this situation, George is the one receiving payment, yet he has not been damaged from the use of personal information. Ken, on the other hand, was affected by the accumulation of this information but remains outside of the equation, as the damages are externalized to him.[150] Ken's information-based transaction was limited to his basic information, which will probably assist the collectors only marginally in creating databases and prediction patterns, yet was used to place Ken within a pattern or cluster created in the past—thus enabling manipulation.

The results of this analysis are problematic. There is no reason to compensate George for the collection of his information, when he does not suffer any harm from its use. Of the two perspectives described above, a transactional-based solution should strive to comply with the latter and compensate individuals for actual damages, rather than for a questionable property right, which was created only to protect from possible abuses of data collection in the first place. These examples establish that regulating the collection stage does not lead to a fair result, but rather provides some collectees with a windfall while others are harmed without recourse through the creation of negative externalities. In other words, transaction-based regulation leads to an unfair (yet random) transfer of wealth from the individuals affected by the use of personal information, to the collectees (who are not necessarily the same individual). This problematic outcome is another result of regulating the personal information market in a phase that is far from the final implementation stage. If the regulation and compensation process occurred at the imple-

150. In other words, the following discussion could be framed using the economics term of externalities. Externalities are best defined as a situation in which the private costs or benefits to the purchasers of goods or services differs from the total social costs or benefits entailed in its production and consumption. An externality exists whenever one individual's actions affect the well being of another individual in ways that need not be paid for according to the existing definition of property rights in the society. For more on the definition of externalities, see the Glossary of Political Economic Terms at http://www.duc.auburn.edu/~johnspm/glossind.html (last visited Oct. 28, 2003).

The external portions of the costs and benefits of producing goodwill will not be factored into its supply and demand functions because rational proti-maximizing buyers and sellers do not take into account costs and benefits they do not actually have to bear. In this instance, the Author's analysis addresses negative externalities from the transaction between George and shoeperstore.com, which adversely affects Ken. Ken, in this example, is negatively impacted by this transaction, while his costs will not be factored into George's demand curve.

The Author has chosen not to frame this analysis by exclusively using the "externality" term, as this Article is not focused on whether the price reaches an equilibrium, but whether the use of a property regime (or the regulation of the "privacy dilemma" in the collection state in general) can achieve a fair result. The Author believes that framing this analysis on the basis of the externality paradigm will also lock us into previous conceptions of market, which might be inappropriate when analyzing transactions based on personal information (which create different and novel problems). For a recent analysis of externalities in the media market, see E. BAKER, MEDIA, MARKETS AND DEMOCRACY (Cambridge University Press, 2002) Chapter 3. For one analysis of externalities with regard to privacy, see Jay Weiser, *Measure of Damages for Violation of Property Rules: Breach of Confidentiality*, 9 U. CHI. L. SCH. ROUNDTABLE 75, 88 (2002).

mentation stage, the "Kens" of this world will be properly protected from or compensated for any problematic uses of personal information while lowering such externalities to a minimum level.

The disparity between these two perspectives and the creation of such externalities can be attributed to data analysis in general and data mining in particular. When data is used in its "raw" form, these two perspectives are basically redundant and lead to the same result. Yet, the sophistication of data mining allows collectors to construct predictive models and make an educated guess about the future conduct of individuals—thus allowing them to externalize. These individuals might have provided only partial information, but fit neatly into patterns previously created using other peoples' data. Therefore, the emergence of data mining technology creates the distinction between collectees and those subject to the use of personal information. As with the earlier discussion regarding the relationship between collectors and collectees, an in-depth analysis of the data mining technology proves that regulating at the collection phase leads to unfair results.

Genetic research is another, more radical, example of a situation in which information about a group of individuals may be transferred to others who share similar traits, or, in other words, negative externalities resulting from the analysis of personal data. When constructing a DNA database, researchers collect specimens and personal information from several individuals who share common traits, such as race, nationality or ancestry. Before such collection, these collectees sign lengthy release forms and may receive compensation for their participation. After collecting and analyzing the genetic data, scientists construct a DNA database, which facilitates their research of the collectees' genetic predisposition to medical conditions and even personal traits.[151] However, the relevance of this information is not limited to the collectees; the DNA database could be used to draw inferences about other members of the larger demographic group. These other members never consented to any part of this genetic research and might have even objected to it. Yet after concluding the initial survey, this does not matter—the genetic data of many of their peers is known to the collectors and therefore strong predictions as to their predispositions are made possible.[152]

The genetic example demonstrates that regulating at the time of collection is insufficient when personal information can be inferred about others that share personal traits with the original collectees. Data mining tools are allowing collectors to draw inferences as well (even though the data mining applications will never provide the accuracy of information available from genetic research). The knowledge derived from the analysis of personal information (both genetic and otherwise) can be used to exploit the public and might not be limited to the specific collectees who consented to the collection of their personal information. Given these difficulties, implementation-based solutions are the only answer.

The above example, from the world of biotechnological ethics, causes great

---

151. *See, e.g.,* GARFINKEL, *supra* note 13, at 190-96. Garfinkel addresses several incidents in this context, such as the mapping of the genome of the entire population of Iceland and research carried out regarding "Jewish Genes." *Id.* The latter revealed specific forms of diseases that occur within the Jewish Ashkenazi population. *Id.*

152. *See* Sally Lehrmen, *Jewish Leaders Seek Genetic Guidance*, 389 NATURE, September 25, 1997, at 322.

concern among scholars, who fear that these practices might fuel bigotry and racism.[153] This Article, however, does not address such dangers, but concentrates on the mundane problems that result from data mining. For instance, vendors may use genetic information in conjunction with the demographic data of their customers to predict their behavior, medical needs and physical vulnerabilities. Thereafter, they might take advantage of their customers through discrimination and manipulation. These issues cause concern with regard to medical insurance,[154] and are sure to spread to other areas.[155]

The basic flaws of the information-based transaction can surely lead to unfair results. Such unfairness pertains both to the relationship between collectors and collectees, as well as the collectees themselves. In view of the critiques presented, the obvious solution would be regulation of both the collection and use of information, thereby eliminating the unfair outcome as well as any fears of surveillance. Yet, such extensive regulation is both unrealistic given the political power of the collectors, and undesirable because of the possible benefits of the collection process.[156] We therefore should strive to come up with an optimal form of regulation that will provide the utmost protection to individuals, while minimally impeding on the collectors' business initiatives and rights. Several implementation-based solutions stroll along this golden path.

III. DOUBLECLICK, INC.—A PRACTICAL PERSPECTIVE

For a realistic look at the problems and solutions in the actual market of personal information, it is helpful to examine DoubleClick, Inc. ("DoubleClick"). DoubleClick represents the first step towards the future practices of information collectors and analysts—the collection of personal information online and its use to distribute advertisements and banners that are tailored to the individual users.[157] These practices open a door to many possible problems: the abuse of personal information by the collectors, possible price discrimination by advertisers and marketers, and even a miniature "autonomy trap" in which the users are receiving advertisements only from one source, perhaps taking advantage of the users' vulnerability at a specific time.[158] Yet, in all fairness, the practices of DoubleClick do not create a real threat to individuals or the public.

---

153. *See* GARFINKEL, *supra* note 13, at 190.

154. *See generally* Makdisi, *supra* note 24.

155. See for example, GARFINKEL, *supra* note 13, at 193-94 (regarding the harm that might befall the Icelandic population).

156. *See* discussion *infra* Part II.

157. According to the Agreement, DoubleClick is defined as a third-party ad service that in turn is defined in section 39 as:

> a business that is a service-provider or vendor to a First-party Web site, not owned or otherwise under the control of that First-party Web site, and whose services may include Online Ad Delivery. For example, during or after a User's visit to a First-party Web site, The First-party Web site may link the User to a Third-party Ad Service, which then transmits Online Ads to the User's computer.

See infra note 161. DoubleClick prefers to refer to itself as "enab[ling] marketers [to] deliver the right message, to the right person, at the right time." DOUBLECLICK, INC., 2000 ANNUAL REPORT 3 (2001).

158. Richard M. Smith, *Internet Privacy: Who Makes the Rules*, Address at YALE SYMP. L. & TECH., ¶ 18 (Spring 2001), a*vailable at* http://lawtech.law.yale.edu/symposium/s01/speech_smith.htm.

DoubleClick's actions certainly did not go unnoticed. They have been the subject of a great deal of public scrutiny and uproar, especially in reaction to DoubleClick's intention to merge its online resources with the immense databases of AbacusDirect, a mass-marketing firm.[159] This contemplated merger, as well as DoubleClick's conduct in general, forced it into the public spotlight. DoubleClick has been subject to inquiry by the Federal Trade Commission,[160] a class action suit[161] and a suit by a group of state Attorneys General (led by the Attorney General of New York). This last suit concluded with a settlement agreement signed on August 26th, 2002.[162] The agreement makes various requirements of DoubleClick, which are spread across the information flow:

A. Collection

DoubleClick rarely collects information directly, but does so through "First Party" websites. In accordance with the Agreement, it is DoubleClick's responsibility that these affiliated websites provide sufficient notice of DoubleClick's practices in their "Privacy Policy."[163] DoubleClick must provide a full explanation of its practices on its website as well,[164] and allow users to "opt out" of DoubleClick's analysis process.[165] The notice requirement, set at the collection phase, is a standard request, yet in view of the mechanics of the information flow, it would have been wiser to require DoubleClick to provide notice and disclosure at the time the personal information is used and implemented, which is when the tailored banner or advertisement is displayed.

In addition, DoubleClick is prohibited from merging its personal information databases with other sources, without the users' (or collectees') permission.[166] This restriction protects individuals who contributed personal information to other databases and were unaware that DoubleClick might use such information in conjunction with its activities. In today's legal setting, it is doubtful whether the states' Attorneys General could have enforced such a harsh remedy via court or legislation, rather than by a settlement agreement.

B. Analysis

First, with regard to the data warehousing stage, the Agreement provides the

159. *See, e.g.*, Greg Miller, *DoubleClick Cancels Plan to Link Net Users' Names, Habits; Internet: Protests Prompt Firm to Halt Project To Combine Databases, Which Could Threaten Web Surfers' Anonymity*, L.A. TIMES, March 3, 2000 at C1; Bob Tedeschi, *In a Shift, DoubleClick Puts Off Its Plan for Wider Use of the Personal Data of Internet Consumers*, N.Y. TIMES, March 3, 2000 at C5; John Schwartz, *Web Firm Halts Profiling Plan; CEO Admits Mistake in Face of Probes, Privacy Complaints*, WASH. POST, March 3, 2000, at A1.

160. *See* FTC letter *available at* http://www.ftc.gov/os/closings/staff/doubleclick.pdf (last visited Sept. 19, 2003).

161. *See In re* DoubleClick, Inc. Privacy Litigation, 154 F. Supp. 2d 497 (S.D.N.Y. 2001).

162. *In re* DoubleClick, Inc., *available at* http://www.oag.state.ny.us/press/2002/aug/aug26a_02_attach.pdf (last visited Dec. 3, 2002) (detailing an agreement between The Attorneys General of the States of Arizona, California, Connecticut, Massachusetts, Michigan, New Jersey, New Mexico, New York, Vermont, and Washington and DoubleClick, Inc.).

163. *Id.* §§ 61-65.

164. *Id.* §§ 58-60.

165. *Id.* § 61(c).

166. *Id.* § 83.

users with a right of access and the ability to delete their profiles.[167]

In addition, the agreement provides restrictions on DoubleClick's data analysis process. The use of "Sensitive Data"[168] in the analysis is prohibited for four years.[169] Such provisions are justified since this form of analysis might lead to problematic forms of discrimination and abuse. However, they are difficult to enforce because they pertain to the inner workings of the firm and will be almost impossible to deduce from observing DoubleClick's business conduct.

DoubleClick is also required to restrict its analysis of additional forms of personal information (defined as "User Data") to the specific uses disclosed at the time of collection (secondary use).[170] This stringent restriction will prohibit new uses of information that might be evident from a data mining analysis (that is not hypothetically-driven). Thus, both collectors and collectees will be unable to benefit from other potentially profitable uses for the collected data.

C. Implementation

This stage, which is usually neglected, surprisingly receives attention in the Agreement. According to the Agreement, DoubleClick is required to develop technology that will allow its users to view any categories associated with their online profile.[171] These categories lead to the selection of the specific advertisement delivered to their browser. This application (the "cookie viewer,")[172] will allow Internet users who receive personally tailored advertisements to obtain information about the way DoubleClick profiles and categorizes them. Such information will theoretically help users become less prone to manipulation by these advertisements, because they have gained some insight into the profiling process. The cookie viewer carries the promise of a technological application that might assist in solving current legal predicaments.

Even though the cookie viewer's concept is indeed a step in the right direction, it is an example of ineffective regulation. The cookie viewer might fit the classic analysis scheme, which is based on niche marketing. However, this appli-

---

167. *Id.* § 80.

168. *Id.* § 36. Section 36 includes the following definition:

> "Sensitive Data" categorically includes but is not limited to data related to an individual's health or medical condition, sexual behavior or orientation, or detailed personal finances, information that appears to relate to children under 13, racial or ethnic origin, political opinions, religious or philosophical opinions or beliefs and trade union membership; PII obtained from individuals who were children under the age of 13 at the time of data collection; and PII otherwise protected under federal law (for example, cable subscriber information or video rental records).

Id.

169. *Id.* § 67.

170. *Id.* § 74.

171. *Id.* § 80. Section 80 states:

> At the time that DoubleClick employs Multi-Site/Multi-Session Ad Serving, DoubleClick will use reasonable efforts to develop technology that allows a User to securely view any Multi-Site/Multi-Session Categories associated with that User's DoubleClick Online Ad Serving Cookie on the User's device.

Id.

172. Press Release, Office of New York State Attorney General Eliot Spitzer, Major Online Advertiser Agrees to Privacy Standards for Online Tracking: Company to Increase Visibility and Verify Data Collection Practices, *available at* http://www.oag.state.ny.us/press/2002/aug/aug26a_02.html (last visited Aug. 26, 2002).

cation is rendered meaningless by the introduction of data-mining tools in the business environment.[173] In this new environment, advanced algorithms replace categories, leaving nothing to view on the cookie viewer.

The DoubleClick Agreement clearly demonstrates how a single issue can be tackled across the various stages of information flow. However, there still remains the issue of how solutions should be formed. However, there are a variety of other solutions rooted in the implementation stage. These are addressed in Part IV.

IV. IMPLEMENTATION-BASED SOLUTIONS

This Article advocates for a specific form of regulation for the personal information conundrum. Such solutions must take into account the possible benefits of information analysis and should preferably be focused on the implementation stage of the data flow. It would be somewhat pretentious to offer an overall solution to all the problems addressed in this Article. Instead, the Author offers general guidelines for some of the problems addressed[174] and a detailed analysis of implementation-based solutions to the possible fears of price discrimination.

A. General

1. Fear of Abuse and Misuse of Personal Information

First, we must tackle the fears of abuse and misuse of personal information. This issue is addressed in a narrow sense, to include publishing information against the will of the individual, or the use of such information to blackmail or embarrass that individual.[175] In general, these matters can be controlled via the legal tools of breach of trust, breach of fiduciary duty and privilege.[176] They are also potentially addressed through various privacy torts, though these are not accepted in all states.[177] Such tools are insufficient to protect the public in the new information age. In today's reality we are providing vast amounts of personal information to those that are not considered fiduciaries, according to present doctrines, and therefore require additional forms of regulations and protections.[178]

To resolve this issue, it has been suggested that government set broader rules regarding trustees and fiduciaries so as to include entities that are currently vested with large quantities of personal data.[179] For example, such doctrines should now include portals, vendors, and credit card companies as trustees as well. These rules will prohibit such newly formed trustees from using personal information in a way that is detrimental to those who entrusted them with such data. A problem-

173. *See generally* Zarsky, *supra* note 103.

174. *See MYOB*, *supra* note 1, at *10-25. *See also supra* note 8 and accompanying text.

175. Beyond the solutions provided below, in extreme situations of misuse, plaintiffs could refer to the tort of intentional infliction of severe emotional distress. *See* Andres v. Bruk, 631 N.Y.S.2d 771 (1995). Though courts tend to dismiss this claim and accept it only in situations of "extreme and outrageous" conduct, it perhaps should be revisited given today's new information landscape.

176. For example, there are several well-known privileges that require the protection of personal information, such as the attorney-client privilege or the doctor-patient privilege.

177. Zimmerman, *supra* note 21.

178. Litman, *supra* note 42, at 1308-09. In the current legal setting, such fiduciary duties or trusts are formed only in the event of an exceptional relationship.

179. *Id.*

atic aspect of this solution is secondary sales, regarding whether they should be permitted and whether the high standards of conduct for trustees and fiduciaries should be expected of the secondary holders of personal information as well.

These broadened standards of trust should apply to some (banks, for instance), but not others (such as vendors). They should also be applied narrowly so as not to inhibit secondary uses for marketing and promotion purposes.

Beyond the use of fiduciaries, courts should consider broadening the use of the tort of "disclosure of private fact" (or "invasion of privacy") referred to above. At this time, the "disclosure" tort is not accepted in all states.[180] Moreover, even where accepted, the plaintiff rarely wins, as the courts require the publicized matter to be highly offensive to a reasonable person, communicated to the public at large and not a legitimate concern of the public.[181] Since the public's standard of "reasonable" is constantly shifting to the extreme, plaintiffs are finding it considerably more difficult to prove the elements of this tort.[182] As Rosen suggests, perhaps courts should set a higher standard of reasonable conduct by which defendants must abide, regardless of society's current subjective standard.[183] However, the "disclosure" tort should still be construed narrowly so as not to include business initiatives that treat customers differently on the basis of their personal information; the detriments that may stem from such business practices are addressed later in this analysis.

These rules should be supplemented by strict regulation of security,[184] by both setting high standards for public companies and entities that accumulate sensitive information, and implementing harsh enforcement against those who breach such security rules.

Obviously, there are additional ways to solve the problems of misuse and abuse. However, the solutions emphasized here concentrate on the final stage of the information flow, in which the actual misuse takes place. At this stage, problems can be confronted successfully, without impeding the possible benefits of information collection.

2. The Adverse Effects of Errors

The second actual harm is the adverse effect of errors in databases.[185] Even though such errors will self-correct as the flow of information continues and intensifies, individuals should be provided with the right to access and correct information in databases, so as to immediately mitigate this problem's present manifestations. Such a right of access will not pose an unmanageable burden on the infor-

180. *See*, Solove *supra* note 4, at 1432.

181. On this issue, see ROSEN, *supra* note 104, at 45-50, 194.

182. For a famous example where a plaintiff failed in his suit against the press arguing this tort, see *Sipple v. Chronicle Publishing Co*., 201 Cal. Rptr. 665, 670 (Cal. App. 1st 1984). For a description of the facts of this case, see ROSEN, *supra* note 104, at 47-48. For another example, see *Sidis v. F-R Publishing Corp*., 113 F.2d 806 (2d Cir. 1940). For a description of the facts of this case, see SMITH, *supra* note 79, at 224-28.

183. ROSEN, *supra* note 104, at 51-53.

184. Security, unlike privacy, refers to unauthorized access to the information. Security is another of the principles of "Fair Information Practices." NAI, *supra* note 3.

185. See ROSEN, *supra* note 104, as to Rosen's debate with Robert Post on whether this issue is really one that is connected to privacy, or is an issue that can arise with regard to public information as well.

mation flow if efficient procedures are put in place, and will not interfere with the collection and analysis stages. As previously mentioned, this right is well established in the literature and practice of information privacy.[186] Furthermore, it is in the best interests of all parties that the databases used for analysis are as accurate as possible. A right of access will surely promote this objective, as well as shield individuals from undesirable results. Individuals' access should be limited, however, to the basic personal information pertaining to every person, rather than to the patterns and categorization carried out by the collectors.

3. Seclusion and Solitude

The third problem society faces from new forms of information collection and analysis is the use of personal information to impose on others' seclusion and solitude. Again, instead of interfering with the collection and analysis stage, this problem can be solved directly through regulation of the implementation stage. For example, several solutions have been suggested and implemented with regard to mass mailing, telemarketing and spam. These solutions allow users to reduce their exposure to problematic practices by obtaining unlisted telephone lines, mandating "do-not-call lists" and opting out of mailing lists.[187] The spamming phenomenon is especially problematic, because the current email infrastructure allows marketers to engage in mass mailings at virtually no cost. In response to this problem, specific legislation has been introduced in several states to protect email accounts from such practices.[188] Technical solutions to these problems are also available and will surely be explored further in the near future.[189]

4. Manipulation

The next problem is the "autonomy trap"—the fear that content providers will use personal information to effectively manipulate the public's choices through the use of personally targeted advertisements and content.[190] These problems concern the providers' ability to affect individuals' views both as consumers (in the context of marketing, for example) and as citizens, with regard to various matters that are on the current national and international agenda.

Implementation-based solutions to these problems require extensive discussion and the Author addresses these issues in another article.[191] In general, however, regulators can tackle any potential attempt to carry out such manipulations by requiring notification and diversity.

By notifying individuals that the content they receive is personally tailored, the individual's awareness will heighten. Notification should detail both the use of personal information in tailoring specific content as well as some indication as

---

186. *See supra* note 78 and accompanying text.

187. *See* Sovern, *supra* note 42, at 1313.

188. *See* Gavin *supra* note 99, at R9.

189. *See* Gleick *supra* note 100, at 42. *See also* John Schwartz, *Consumers Finding Ways to Zap Telemarketer Calls*, N.Y. TIMES, Dec. 18, 2002.

190. For the definition of the autonomy trap, see *supra* note 95.

191. *See* Tal Zarsky, *Disarming the Trap; Defining, Explaining and Solving the Challenges to Autonomy through the Use of Personal Information Analysis in the Internet Society* (Working paper, on file with Author).

to how this information is used. With this knowledge, users and information recipients will be reminded to use their judgment when assessing and evaluating the information they receive.

Ensuring that individuals receive a variety of advertisements from various sources would offset the specific negative effects that any single source of tailored content might have. Thus, to defeat manipulation, access to and use of several content providers or forms of content must be secured and even encouraged.

Both solutions pose several problems. Their introduction forms an impediment on the content providers' free speech rights, thus requiring further constitutional analysis. Moreover, the effectiveness of such solutions is questionable and needs to be explored within various scenarios. The Author takes on these challenges elsewhere.[192]

B. Price Discrimination

Finally, there is a possibility that personal information will be used to facilitate price discrimination between clients and customers. This issue includes several sub-topics with various levels of severity.

1. General

At this time, vendors and marketers can make use of personal information they obtain to create different pricing schemes for different types of customers. The Internet and E-commerce environment provide fertile ground for such practices, since vendors can easily collect personal information about every user, and create a different "store" for every customer by providing them with a different screen or window. This way, the customer does not know he or she is receiving service and treatment that is different from others and will not suspect being overcharged.[193]

Even though such practices may seem unfair at first sight, they are essential to achieve fairness in pricing and avoid re-distribution within non-homogonous groups.[194] The problematic aspects of price discrimination become evident only when viewing the underlying factors that are the basis for such discrimination.

As mentioned elsewhere,[195] such dynamics may seem unfair, but they are usually resolved naturally by the forces of the market. There are several instances in which the state should consider welfare interests to ensure fair treatment of the underprivileged. Two examples of such instances are discussed below: insurance and credit. Beyond these situations, the state should not interfere with this form of dynamic pricing.

---

192. *See generally id.*

193. See *MYOB*, *supra* note 1 at nn.59-65 and accompanying text.

194. The situation in which every customer is provided with a price that is exactly tailored for her is referred to as "perfect price discrimination." For the benefits of such practices, see Jonathan Weinberg, *Hardware Based ID, Rights Management, and Trusted Systems*, 52 STAN. L. REV. 1251, 1274 (2000). In addition to the benefits mentioned, price discrimination is desirable since it broadens the distribution of goods, especially with regard to informational goods such as copyright. *See* William W. Fisher III, *Property and Contract on the Internet*, 73 CHI-KENT L. REV. 1203, 1239 (1998).

195. *MYOB*, *supra* note 1 at nn.66-69 and accompanying text.

Price discrimination can be carried out on the basis of race, nationality and sexual orientation.[196] Pricing can be motivated by bigotry, or simply as an attempt to seek profits and increase revenues. As these forms of discrimination are clearly contrary to public policy, they require direct intervention and regulation, and should be categorically prohibited. This requirement is already implemented in various ways, either by prohibiting the collection of such factors, or prohibiting analysis on the basis of this information.[197] Even though these steps interfere directly with the information flow, the results of such discriminatory practices and analysis are harmful to society and should therefore be avoided.

As the Author explains elsewhere,[198] sellers and marketers might use the personal information they collect to facilitate price discrimination between clients while taking advantage of their vulnerabilities or lack of knowledge. By using their access to personal information, vendors can accurately assess the highest price the buyer will be willing to pay and thus minimize their risk in a transaction. Vendors can overcharge when the data collected indicates that the buyer is indifferent, uninformed or in a hurry. The buyers, on the other hand, do not have these capabilities and are not privy to similar information about the sellers. They are unable to carry out such analysis regarding the collectors and are therefore at a considerable disadvantage. These practices are especially suitable in the Internet setting, where the customer provides a vast amount of personal information and can be provided with a personally tailored "market for one." These problems can be solved directly without affecting the collection and analysis stages.

One simple solution is to prohibit these practices when the marked-up prices rise above a certain level.[199] Recent scholarship discusses the applicability of several legal doctrines to enact such rules.[200] Initiating this type of regulatory scheme will introduce several challenges because of interference with transactions between consenting parties in the open market. In this context, antitrust doctrines seem most relevant, but the dangers of price discrimination are not limited to the actions of monopolies as normally defined. A broader and more flexible definition of an affective monopoly will be required in order to indicate in which instances inflated prices could be subject to scrutiny and which sellers must abide by these

196. See *MYOB*, *supra* note 1 at n.74 and accompanying text. *See also* Ian Ayres, *Further Evidence of Discrimination in New Car Negotiations and Estimates of Its Cause*, 94 MICH. L. REV. 109, 137 (1995) (explaining why sellers might use such factors as a basis for discrimination for reasons other than bigotry).

197. As noted, such steps have been taken in the DoubleClick example provided above. *Supra* Part II. Such steps have also been introduced as part of the EU Directive. *Supra* note 93 and accompanying text (discussing the analysis of sensitive information).

198. *MYOB*, *supra* note 1, at *12-17.

199. The most obvious law to regulate price discrimination is the Robinson Patman Act, 15 U.S.C. § 13(a) (1982), which prohibits price discrimination explicitly. The elements of a prima facie case include: (1) a price difference; (2) between sales to two buyers; (3) of commodities; (4) of like grade and quality; (5) that creates a reasonable possibility or probability of competitive injury.

However, today's academic discussion mostly regards this Act as irrelevant and unusable. *See* Mark Klock, *Unconscionability and Price Discrimination*, 69 TENN. L. REV. 317, 359-60, 370 (2002). *See also MYOB*, *supra* note 1, at *12 n.64. However, it is possible that this Act will be used as a foundation for future antitrust regulations regarding this issue.

200. *See, e.g.*, Klock, *supra* note 199 at 359-60, 370.

201. At the time of the transaction, the buyer might have other options for purchase. Yet for various reasons he cannot or does not make use of them. Therefore, the seller has an affective monopoly at that specific juncture.

new rules.[201]

Instead of constructing cumbersome regulatory schemes, price discrimination can be confronted by encouraging market forces to oppose such pricing dynamics. Scholars specify three conditions that must be fulfilled for price discrimination to occur: (1) the indicated price must be within the consumer's range, (2) the seller has market power, and (3) there is no secondary market (arbitrage).[202] This third condition is where the sellers are most vulnerable.

The key to undermining the price discrimination practices is in promoting secondary markets. We need not go so far as to facilitate a secondary market for the actual products—creating a flow of information about the products would be sufficient. If customers can communicate with others in the same market (or better yet, at the same store), they could receive strong indications as to whether they are overcharged in any given transaction (information that is not easily obtained in the Internet setting). With this line of communication in place, consumers could transact among themselves to avoid being charged the marked-up price.

Solutions along these lines are indeed possible yet do not require "standard" regulatory proposals; instead, changes in the infrastructure (or code) of e-commerce or marketing websites are required. The key to such solutions is mandating or encouraging the installation and use of communication channels between users visiting a specific website. Several software tools could be applied to fulfill this requirement, such as: "chat rooms" that will be affiliated to the e-commerce web site yet independent of the vendor,[203] where customers could interact with others during their "visit" to the relevant website; the use of instant messages within the site; or tools that will enable customers to leave notes or comments at every virtual juncture for future visitors.

With these tools, communication and interactions between users could be carried out in "real-time" and facilitate those customers that are locked into higher prices due to time or information constraints. Using such means of communications, customers could inform each other of better bargains with other vendors, or perhaps even with the same vendor. This flow of information will facilitate a secondary market, which will discourage vendors from attempting to carry out price discrimination schemes. It would also provide a partial response to a vendors' attempt to provide different segments of the public with customized products, which might be different in design and price. With this form of communication, customers could interact with ease, compare the specifications of the various products and opt for the product with the highest value for its price.

There are two caveats to this technology-based solution to the arbitrage (and therefore the larger price discrimination) problem. First, providing *too much* information to purchasers may have adverse effects as well. Providing buyers with a "full picture" of all the prices and pricing schemes used by the sellers will tilt the information and risk imbalance towards the buyers, to the clear detriment of the vendors. In addition, it will make any form of price discrimination (which, as mentioned above, has positive aspects and uses as well) almost impossible. However, the described solution does not provide buyers with complete information,

202. Weinberg, *supra* note 194, at 1274.

203. Such independence is essential so that the sellers cannot interfere with these communications.

but only with an opportunity to bridge the information gap, and use diligence to try to obtain a better deal, thereby striking a fair informational balance between sellers and buyers.

Second, as with any solution that is based on "code," the ideas and concepts introduced may seem appropriate to the Internet environment, but are more problematic when applied to the "brick and mortar" world where "chat rooms and instant messages cannot be linked to the vendor's place of business."[204] Therefore, other forms of regulation, such as applying antitrust doctrines to "market up" prices and affective monopolies, should be implemented for the offline setting. Note however, that as opposed to collection that is widely practiced both on and off line, the one-on-one marketing opportunities are not as popular or possible in a physical environment. In the physical world there are more opportunities for interactions between customers, which will allow for the flow of information among them. This facilitates arbitrage markets even without the active encouragement of a regulator.

2. Exceptions

As with every rule, this analysis of price discrimination must confront several exceptions that do not fit neatly within the paradigms presented and therefore require additional discussion. Consider the following examples:

a. Medical and Life Insurance

The price of insurance to cover unfortunate occurrences such as disability or death of a family member (life insurance), or any future medical treatment (medical insurance) relies heavily on the analysis of personal information. Unlike the situations described above, in the insurance example, personal information is not used to detect the consumers "demand curve" or the highest price they are willing to pay for this service. Rather, the personal information is used for actuary calculations to assess the chance that the events a policy covers will occur. Clearly, insurance companies use the results of such actuary calculations to set the premiums on every insurance policy. In this context, it is difficult to label the price differentiation between various insurance policies as price discrimination, since the "products" the customers receive are not the same. Various policies are priced differently, yet they provide coverage for events that might occur at varying levels of probability.[205] The insurance example also does not refer to practices that are usually carried out online, or take advantage of the one-on-one marketing schemes the Internet creates so successfully. Nevertheless, this example presents a specific aspect of problems arising from the advance means of collection and analysis of personal data.

As a result, some individuals are required to pay a very high premium for insurance. The insurance companies might charge these high premiums when the relevant personal information, as well as the patterns and correlations derived from it, indicate a high-risk customer. High fees might be out of these individuals'

204. This issue parallels the Author's discussion regarding P3P and its inapplicability to collection carried out in the physical world. See *supra* notes 47-50 and accompanying text.

205. Price discrimination is best defined as providing the same product or service to different people for different prices. In this context, this definition does not describe the actual process.

financial reach and considerably higher than the rates of others. The firms might charge high prices even when the relevant persons are in perfect health, but their personal information indicates a tendency towards events that will result in additional expenses to the insurance company.

Even though the market for insurance policies is not yet dominated by monopolist players, competition will not resolve these problems. Since the relevant information might be available to all insurers,[206] the prices set for the risky candidate by the competing insurers will be equally high. In addition, individuals will find it extremely difficult to weigh one policy against the other given the fact that the services provided are different for every policy package (which encumbers the dynamics of a competitive market).[207] Even if competitive markets lead to lower rates, such a process will take time. Meanwhile, many people will be faced with the grim reality in which they cannot afford the medical or life insurance they require.

Why is this result problematic? From the insurer's perspective, the abovementioned analysis of personal information enhances their efficiency, reduces their losses on risky policies, increases the profit of their shareholders, and allows them to lower the policy rates for the general population. The advanced analysis of additional personal information assists in regrouping the policyholders, so that those who present lower risks will not pay the higher rates that others, who do present higher risks, are charged.

However, the personalization of the premium price creates a social problem, since the notion of "insurance" can be understood on two different levels. On the most basic and practical level, insurance allows individuals to spread the risk of detrimental occurrences over several years and share it with other people. Insurance also represents the notion that it is both fair and economically efficient that no single individual should be confronted alone with the burden of personal tragedies. According to this latter concept, insurance rates should be set at an affordable level, so that all members of society can share the risk.[208] When more and more people are marked as "high risk," they are placed outside this supportive circle and must face the grim reality alone. Thus, the use of personal information in the assessment of risk defeats this social goal.

The question currently debated by legislators and academia is the extent to which personal information (which now includes genetic data as well) can be used in this context given the two goals of insurance.[209] There are several ways to settle the tension between these two concepts. The basic solution, which to a cer-

---

206. This is accomplished by means of the secondary market for personal information.

207. This is actually a "transactions costs" claim (variations of which were presented above regarding the information based transaction in general). *See* discussion *supra* Part I.

208. JON S. HANSON, REGULATION OF THE LIFE INSURANCE BUSINESS 17-20 (1996) (providing the two goals of insurance regulation—internal and external). The external regulations focus on the socialization of risk and address the issues of availability of coverage and reasonability of price. *Id*.

209. Among other issues, today's debates focus on two difficult questions. First, the degree to which personal genetic information may be used. *See* Eric Mills Holmes, *Solving the Insurance/Genetic Fair/Unfair Discrimination Dilemma in Light of the Human Genome Project*, 85 KY. L.J. 503 (1996-97). Second, whether information regarding HIV could be used when calculating an insurance premium. *See* HANSON, *supra* note 208, at 156-67; Alan I. Widiss, *To Insure or Not to Insure Persons Infected with the Virus That Causes AIDS*, 77 IOWA L. REV. 1617 (1992).

tain degree is currently implemented, requires blocking the collection and use of personal information—a "collection-based" solution.[210] However, this solution is not preferable. Beyond the notion of privacy that this Article does not discuss[211] and the benefits the analysis of personal information can bring about, these solutions are simply unfair.

Up until now, the unavailability of our personal information caused a socially desirable result. The "fog" as to what the future has in store required insurers to treat most policyholders equally by setting standard rates.[212] These rates consisted of a subsidy of those policyholders who are a higher risk to the insurance company. By prohibiting the analysis of personal information in this context, the "fog" will not clear out and this subsidy will remain. Many of the elements that form this subsidy are arbitrary. In today's world, even though some individuals are protected from higher insurance rates by the regulation of information collection, others are not. There are always those instances when insurance companies lawfully obtain the individual's personal information, and later use it to classify an individual as high risk. For example, such information could be made available to the firm through its mandatory questionnaires, or from bills it pays for medical services. Therefore, blocking the collection and analysis of information might benefit some, but not all customers. The firm will always have the ability to obtain information about certain clients, while remaining unaware of the traits of others; the differentiating factors between these groups are nothing short of chance.[213]

In view of this flaw, society should not pursue its social goals while relying on the partial opacity of personal information, but confront the matter directly in the implementation stage. Instead of solving the problems of insurance by placing impediments on data collection, government should set maximum insurance rates or require that specific factors not be taken into consideration (such as gender or handicap),[214] and offer subsidized rates or "no fault" insurance.[215] This way, society could preserve the social advantages of insurance, avoid the arbitrary results mentioned and at the same time reap the benefits of information collection.

---

210. This is today's regulatory trend with regard to genetic information. Holmes, *supra* note 209, at 662.

211. *See* HANSON, *supra* note 208, at 160 (suggesting that even though the privacy issue is addressed often, it is not central to this discussion, as most of those that obtain personal information are bound by confidentiality rules).

212. Holmes, *supra* note 209, at 577 (discussing Rawlsian Theory and the way we live behind a "veil of ignorance" with regard to our future health).

213. To avoid this result, there is always the option to ban all uses of personal information. However, as regulators are reluctant to restrict the use of information within the firm, *see* discussion *supra* Part I, and such harsh regulation does not seem realistic at this time.

214. *See generally*, Ariz. Governing Comm. for Tax Deferred Annuity & Deferred Comp. Plans v. Norris, 463 U.S. 1073, 1086 (1983) (finding that the fact that gender caused a disparity in the annuity received was unlawful); L.A. Dep't of Water & Power v. Manhart, 435 U.S. 702, 717 (1978) (holding that disparity in the premium paid based on gender is unlawful). However, these are Title VII cases that refer to discrimination in employment. Several states have enacted rules regarding the issue of discrimination on the basis of gender and other problematic factors with regard to insurance rates. *See* HANSON, *supra* note 208, at 130-31. Similar problems arose regarding the calculation of risk for the handicapped. *Id*. at 131-32.

215. *See* Holmes, *supra* note 209, at 661.

b. Credit Rating

Personal information also plays a dominant role in setting credit ratings and interest rates for individuals seeking financing. Here, personal information is used to tailor the rate the lender charges the borrower. Again, such practices do not stem from the "one-on-one" environment, and do not appear dominant in the Internet setting.[216]

The analysis of personal information can be used in a similar manner as that referred to above with regard to insurance—calculating risk and predicting customers' demand. At first, lenders use personal information to broaden their knowledge regarding the relevant borrower's "demand" or their "reserve price." Attempts to discriminate on the basis of such knowledge will be mitigated by competition in the open market. Moreover, those seeking financing can also search for financial institutions that are in need of clients and who offer the lowest rates, thus leveling the information playing field and promoting competition. However, personal information can also be used to analyze the underlying risk of default in any debt transaction. Firms use personal information to set personally tailored interest rates for every client, based on the calculated risk of default. Again, it is difficult to refer to such analysis and pricing schemes as price discrimination, since every loan is a different product and carries a different level of risk.

The same rationales addressed in the insurance analysis are, to a certain extent, applicable here as well. On the one hand, lenders aim to efficiently differentiate between borrowers that present varying levels of risk, and act in the interest of their shareholders and other customers. On the other, the use of personal information in such analysis creates a reality in which certain individuals (classified as high risk due to specific personal attributes) are unable to obtain loans and mortgages, since they face interest rates outside of their financial reach.

These problems should not be solved by prohibiting the collection of information and thus maintaining an "opaque" reality (the collection stage solution) due to potentially arbitrary results. Here, government should refrain from interfering with the market by setting maximum rates, but should protect social interests by providing government-backed loans at comfortable levels for specific essential objectives (a solution that is already widely implemented).[217]

V. CONCLUSION

Solving the personal information conundrum is not a task for the weary. This Article suggests a path through the obstacles that any attempt to regulate these issues will surely face. It does so on the basis of the information flow paradigm, introducing the effects of data mining analysis, and the growth of the Internet society. Clearly, several issues still remain unsolved. However, this Article suggests several possible solutions, which require that we keep in mind both the benefits of

216. The Author's opinion regarding this point may change, given that several companies are trying to promote the use of the Internet for these practices. Many of such ventures try to advertise the attractiveness of the Internet for these transactions by stressing the fact that in this environment the client could have access to various offers and choose the best one with ease. Therefore, this market is still taking shape and it is unclear what it will resemble in a few years.

217. One example is the federal student loan program.

the information analysis and fairness among all the participants in the information market.

As the realm of surveillance grows and the ability to analyze data is enhanced, the accepted notion of privacy is bound to change. It is up to legislators, and to a greater extent, legal scholars, to ensure that new boundaries allow individuals to be protected from abuse, yet still benefit from the advantages the information can offer. The best way to ensure such protection is by enacting implementation-based rules.