Communications of the Association for Information Systems

Volume 37 Article 34

10-2015

Data Mining and Privacy: An Initial Attempt at a Comprehensive Code of Conduct for Online Business

Dinah Payne
University of New Orleans, dmpayne@uno.edu

Brett J. L. Landry
University of Dallas

Matthew D. Dean
University of Southern Maine

Recommended Citation
Payne, Dinah; Landry, Brett J. L.; and Dean, Matthew D. (2015) "Data Mining and Privacy: An Initial Attempt at a Comprehensive Code of Conduct for Online Business," *Communications of the Association for Information Systems*: Vol. 37 , Article 34.
DOI: 10.17705/1CAIS.03734
Available at: https://aisel.aisnet.org/cais/vol37/iss1/34

Communications of the Association for Information Systems

Research Paper ISSN: 1529-3181

Data Mining and Privacy: An Initial Attempt at a Comprehensive Code of Conduct for Online Business

Dinah Payne
Department of Management, University of New Orleans
dmpayne@uno.edu

Brett J. L. Landry
Satish & Yasmin Gupta College of Business, University of Dallas

Matthew D. Dean
School of Business, University of Southern Maine

Abstract:

The prevalence of data mining by businesses and government organizations raises concerns among many individuals about the privacy of their personal data. We address this issue by offering a different perspective that reconciles the conflicting desires of businesses and consumers. We describe privacy, data mining, and their interaction in the larger context, identify the costs and benefits of the uses of data mining, and discuss potential stakeholders found at the intersection of the two subjects. To help synthesize our proposed code of ethical conduct, we examine existing codes of conduct and how they relate to the issue of privacy in the context of data mining with people, processes, and technology. Showing that a uniform code of ethical conduct for online privacy is feasible from both a managerial and ethical perspective, we provide an initial philosophical and principle synthesis that businesses and organizations can tailor for their own specific customers and needs. The developed code of ethical conduct respects consumers' desire for privacy while allowing businesses to use data mining techniques to elicit information that benefits both the business and the consumer.

Keywords: Data Mining, Privacy, Ethics, Learning, Knowledge, Code of Conduct.

This manuscript underwent peer review. It was received 10/15/2013 and was with the authors 6 months for 3 revisions. The Associate Editor chose to remain anonymous.

1 Introduction

In this paper, we review some facts related to the usage of online data mining techniques and the reasons for engaging in data mining to explore the substance of the potential benefits and costs associated with data mining. For example, Hoffman (2012) suggests that many cities in the United States (US) are releasing their data sets to the public, as have cities and governments around the world. Purposes for which governments are data mining include engaging members of the public in efforts to track their performance and keeping abreast of legislative intent and execution. Other salient facts related to the ubiquity of data mining relate to concerns individuals have that their information is not being kept private or properly secured: Kucan (2009) reports that two-thirds of the citizens in the European Union (EU) are concerned about privacy and security of personal information; in Austria and Germany, 90 percent expressed such concern.

The issues regarding data privacy and data mining are further complicated by a changing landscape of technology and tracking. Tracking includes not only online behaviors and history but also physical tracking of an individual with mobile and GPS devices. Never before the current era could an individual be tracked and their data logged by so many devices that data mining efforts can use. In this paper, we address some of those issues and offer a different perspective to the business community about how to reconcile their desire to use data mining for legitimate purposes that consumers favor to maintain some level of privacy.

This paper proceeds as follows: in Section 1, we identify the critical vocabulary associated with data mining and provide background on their definitions, some of which are still entirely debatable. In Section 2, we identify the benefits and costs associated with data mining and some of its uses. Further, we discuss potential stakeholders. In Section 3, we examine characteristics of good codes of conduct not only broadly as they relate to all sorts of moral business dilemmas but also more specifically as they relate to the issue of privacy in data mining. To do so, we review a survey of extant codes of conduct both domestically in the US and abroad. Finally, in Section 4, we develop a code of ethical conduct businesses can adopt that will respect consumers' wishes to maintain some level of privacy while also allowing businesses to elicit information that will help them better serve consumers and, thereby, our economy

2 Basic Definitions Reveal Complexity of Online Privacy Issues

To begin, we define several concepts before debating solutions to the conundrum of business desire to obtain information and consumer desire to maintain privacy. These concepts include privacy, right to privacy, data, personal data, information, data protection. Moreover, we define data mining itself. Meta-ethics suggests that, in any ethical or even practical debate, it is critically important for the debaters to work from the same frameworks: thus, we should come to some agreement on what the basic ideals are before we can generate a code of ethics that would be embraceable by all.

De George (2010), noting that the concept of information privacy is vague and controversial, defines it as a "claimed right on the part of individuals to keep information about themselves private" (p. 466). van Wel & Royakkers, 2004, p. 131) state that "Informational privacy mainly concerns the control of information about oneself. It refers to the ability of the individual to protect information about himself". Tsai et al. (2011) note that privacy is very hard to define with clarity. Kucan (2009) distinguishes privacy and data protection: privacy is a negative right not to be interfered with, while data protection is a positive right held by consumers that their information is collected using certain protocols to ensure appropriate levels of privacy and security. Additionally, the knowledge inside the organization may be explicit in that it is clearly defined, documented, and integrated into policy and procedures and tacit that is based on experiences and intuition (Koh, Gunasekaran, Thomas, & Arunachalam, 2005).

The right to privacy is also a concept that needs review. Collingwood (2012) notes that, in the United Kingdom, there is no legally recognized right to privacy, though British citizens have a general right to respect for private and family life. Further, she notes that judges undertake a "rigorous balancing exercise" (p. 328) in grappling with the right to privacy and freedom of expression. Milne and Gordon (1993) use the idea of implied social contract in discussion on rights to privacy: consumers have the right to proper treatment of their private information. Fule and Roddick (2004) impliedly concede that there is not an absolute right to privacy because even our existence in society requires interaction that would necessitate providing some personal or private information to others.

Legally, researchers have defined personal data as that automatically processed by equipment (Azmi, 2011). Personal information has been presented as five different things including privacy sensitive information, which may be personally identifiable information (PII), sensitive information, which requires more safeguards, and sensitive PII culled available in the public domain. Usage data and unique device identities, which are directly traceable to the individual, are also types of personal information (Pearson, 2009). Table 1 presents more specific information on these kinds of information.

Table 1. Types of Personal Information (Adapted from Pearson, 2009)

Personal information	Examples	Optimal security level
Personally identifiable information	Information used to identify or locate an individual, such as name or address, or information that could be correlated with other information to identify someone	Moderate
Sensitive information	Information related to religion, race, health, sexual orientation, personal financial information, etc.	High
Sensitive personally identifiable information	Sensitive PII information such as biometric information or information culled from surveillance cameras in public places	Moderate
Usage data	Information associated with the use of a computer, smart phone, or mobile device such as web browsing history, application logging, and geotagging to track an individual	High
Unique device identities	Other kinds of information that may be uniquely traceable to a user device	Moderate

Data mining is a technique of using special software to filter large databases to derive information that is implicit rather than explicit (De George, 2010). Knowledge discovery in databases (KDD) is the overall process of discovery of knowledge one can use for a variety of purposes; thus, data mining is merely a part of the KDD process. KDD comprises five steps: data collection and cleansing, choice of pattern discovery method, pattern discovery, pattern presentation, and use of the knowledge discovered (Mehta & Dang, 2011). Information provided for one purpose can be used to determine relevant correlations such that consumers or potential consumers can be better served with more individual attention and/or identified or targeted. Jackson (2002) defines data mining as having the objective of identifying new, possibly useful and meaningful correlations and patterns in existing data. Jain, Yadav, and Panday (2011) add to data mining's definition the idea that it is data from different points of view that aid in one's creating useful information. Further, they identify a large number of applications for mined information: advertising and marketing efforts, bioinformatics, fraud detection, e-commerce, health care, security, financial forecasting, and so on.

2.1 Benefits and Costs of Data Mining

Danna and Gandy (2002) describe several benefits that organizations and consumers can derive from data mining. First, using data mining in customer relationship management (CRM) systems helps retain customers, and it is cheaper to retain customers than to continuously seek new ones. Second, by engaging in data mining, an organization can increase its market share by customizing itself to adapt to customers' general and specific needs and to provide a better "360°" view of customers. Third, data mining techniques' ease of use makes serving customers better and easier: a "no-brainer" reason to engage in it. Fourth, using data mining techniques also allows organizations to more frequently respond in real time (even instantaneously) to consumers' enquiries. Finally, data mining return-on-investment measurements indicate that, over the lifetime of the consumer, customization made possible by data mining has made a good return on the firm's investment of time and money. In fact, Tsai, Egelman, Cranor, and Acquist (2011) empirically found that consumers would pay a premium to keep information private, which suggests that even consumers think that divulging private information is valuable.

Smith, Milberg, and Burke (1996) have identified concerns that consumers are worried about four things: information collection, unauthorized secondary use of collected information, information mistakes, and other improper access to personal information. In addition to access controls, knowledge management systems (KMS) contain structured and unstructured systems that require different security concerns than traditional data and information systems and may not be adequately protected (Randeree, 2011). Malhotra, Kim, and Agarwal (2004) found that the collection of personal information, control over the

collected information, and awareness of how the information would be used are also concerns that consumers felt regarding their personal information. Other costs associated with data mining include difficulties associated with consumers' ability to effectively understand what is in privacy policies; indeed, just to read one for each Internet interaction that might have such a policy is onerous in the extreme, to say nothing of the fact that the "jargon" used is beyond many consumers' comprehension (Cranor, 2005; Kelley, 2009; Jensen & Potts, 2003; Tsai et al., 2011). For an exhaustive list of data mining's benefits and costs, see Cook and Cook (2003); similar to some of the perspectives we outline below, they divide the benefits and drawbacks into categories related to business, individuals, and society.

2.2 Types and Uses of Mined Information

Cranor (2003) identifies four types of personalization systems, all of which address different kinds of information: explicit vs. implicit data collection, duration, user involvement, and reliance on predictions. Personalization, the use of private information about consumers, can be done with regard to demographic characteristics, preferences, and ratings. Search queries, purchase history, and browsing history are also kinds of information consumers provide either intentionally or unintentionally that businesses can find useful in creating profiles of customers. Using any of these methods for developing consumer profiles, the business may then more specifically target consumers and match them with the business's provision of goods and/or services. Data mining activities can be grouped into four areas: efficiency, security, customer-service, and innovation (Payne & Landry, 2012), and all can be used in organizations regardless of sector. What is important is to identify the stakeholders affected by the collected data and how it is processed.

2.3 The Perspectives: Stakeholders and Issues

Stakeholders are anyone that could be positively or negatively affected by some action (Raiborn & Payne, 1990). Payne and Trumbach (2009) identify several stakeholders in a study of data mining ethics: customers/clients of data warehouses, data warehouse management, subjects of information searches, society at large, professional associations with oversight interest in data mining, governmental regulators, competitors, information suppliers, and current and future financial supporters of the firm. Baumer, Poindexter, and Earp (2004) are more specific in citing the Department of Defense, the Department of Homeland Security, the Federal Trade Commission, and the Transportation Security Administration: all these agencies have some oversight over and interest in the collection of personal information. Figure 1 represents all of the stakeholders listed above.

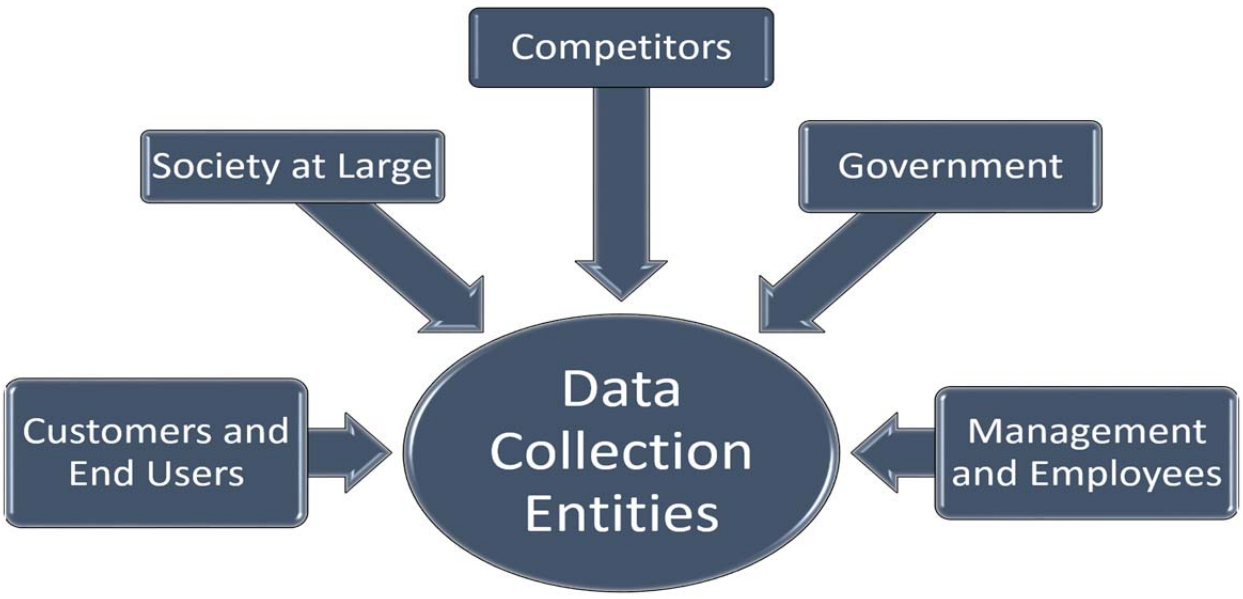

Figure 1. Stakeholders in the Data Mining Privacy Debate

2.4 Stakeholder Obligations: Examples

Business is in business to make a profit: as De George (2010) notes, business engages in all the business functions to serve the rational end of making a profit and providing value to the consuming public. It is management's responsibility to make this happen effectively and efficiently to the betterment of all members of society. The tools that data mining provides (e.g., better targeting and tailoring of their business to users' needs and wants) are invaluable to management in its profit-making endeavors. Further, management and a firm's employees both have corresponding rights and responsibilities: as they have expectations of privacy for themselves, personally and professionally, so does management have the obligation to alert employees about the privacy-sensitive nature of certain consumer information and employees to respect such information (De George, 2010). Thus, management and the firm's employees as stakeholders have the right to make a profit from their business activities and an obligation to do so in an ethical way, as they themselves would want to be treated.

There is another stakeholder, the customer, who is critically important in any discourse on data mining. They provide the information while knowing it could be used—properly or improperly. A resolution to this whole issue could be brought about by a change in the data subject's attitude: the consumer should simply realize that there is no online privacy and act accordingly (De George, 2010; van Wel & Royakkers, 2004). However, this solution is facile and not likely to meet consumers', management's, or any number of other stakeholders' goals. In a similar vein, Fule and Roddick (2004) suggest that unrestricted data mining is a possible choice as long as firms alert consumers that there will be unrestricted data mining (i.e., managing risk rather than eliminating it). Jensen and Potts (2003, 2004) suggest that consumers have responsibilities of their own: to inform themselves as to privacy policies and to act accordingly. These responsibilities are congruent with the idea that everyone is responsible for their own behavior: if we abdicate our rights to make informed decisions, we have assisted in the potential violations of privacy data mining can create.

3 More Issues: A Comprehensive Study on Data Mining Policies

In a comprehensive comparative study of European and U.S. corporate privacy practices, the Ponemon Institute (2006) developed information in nine substantive areas. The comparison relates to the existence and content over privacy policies and their uses. Communication and training questions concerned the ease of understanding of the policy, the process for communicating the policy to those affected, and training policies for the firm's data handlers. Another comparison relates to privacy management: the questions here relate to management's role in establishing and implementing privacy policies. Data security methods were studied, too: technologies used to secure data, privacy platforms, and cookies (text based tracking files on the user's computer) use were surveyed. The sixth and seventh issues are privacy compliance, and choice and consent. The Institute also reviews cross-national standard examination among businesses to see if the international element of data mining was of import to the firms surveyed. Finally, the consumer's ability to redress problems encountered with privacy policies is presented. The study also looks at more procedural issues: it reviews budgets given to firms to establish, implement, control privacy policy and issues, and examines the maturity of privacy policy efforts (i.e., has the firm just begun to examine privacy is a data mining issue for themselves and their consumers or have they been managing a privacy policy for some time). These elements studied provide insight as to the concerns felt by business, consumers, and society: drawing on all these perspectives, the authors of that study develop a "laundry list" of good questions any business, consumer, or regulatory body could ask to ascertain a firm's commitment to preserving data privacy.

3.1 Elements of a Good Privacy Policy

> *There is a manifest need for the negotiation of an international, technology-neutral, certifiable, management standard for the implementation of the information privacy principles that may be implemented by any public or private organization that collects, uses, process or discloses personal information via the Internet, or through any other public or private network.* (Bennett, 2000, p. 33)

Bennett (2000) cites the above as a contemporary reality about privacy protection in our global economy. Policies or codes of conduct should be clear, comprehensive, positive, and enforceable (Raiborn & Payne, 1990). Policies should be clear and attempt to reduce or eliminate ambiguity or doubt in the ethical decision making process. Policies should be comprehensive in that they should be effective guides

regardless of the moral dilemma in which decision makers find themselves. They should be positive in nature by expecting that business professionals in particular will do the right thing and do not need be cautioned or threatened if they should do the wrong thing.

Lastly, the policy should be enforceable: as De George (2010) notes, failure to enforce codes of ethics renders them useless (Raiborn & Payne, 1990). Cook and Cook (2003, p. 397) suggest a great reason for accountability:

> *following ethical practices and respecting the privacy of individuals makes good business sense. Bad publicity associated with a single incident can taint a company's reputation for years, even when that company has followed the law and done everything that it perceives possible to ensure the privacy of those from whom the data was gathered*.

Wilder and Soat (2001) report that "ethical, privacy-respecting practices simply make good business sense" (p. 2).

In our efforts to fashion a comprehensive code of ethical conduct for privacy in data mining, we reviewed several existing policies. These include legal and voluntary codes established by various governments and organizations. These codes provide the basic elements of our comprehensive code of ethics for data mining privacy and integrity.

3.1.1 ISTPA Framework

The International Security, Trust, and Privacy Alliance (ISTPA) (2001) framework project presents one set of privacy principles one can use to develop our uniform code of ethics for online privacy policies. They list the principles as disclosure notice, relevance of information used, participation by the data subject, collection and use limitations on the kinds of information to be mined and used, accountability of the miner(s), security preservation of the information (data confidentiality), and verification of the information (data integrity). The ISTPA also identify practices that should be used in data mining to implement the suggested privacy principles: notice and awareness to/of the data subject, choice, consent and individual access of said subject to the data mining process and results, information quality and integrity checks (including the ability to update and correct misinformation), and enforcement of privacy policies and recourse against transgressors of the policies.

3.1.2 EU Directives

The EU has also issued Directives in the area of data mining. It is illegal to archive research data without the data collector or storage entity having, prior to the collection and/or storage of the data, been given permission to engage in those activities (Carusi & Jirotka, 2009). Notably, Carusi and Jirotka discuss three major E.U. principles of good codes of conduct relating to online privacy policies that are incorporated into all the codes we found. They include making sure that there are informed consent options, anonymization of information when desired by both the data subject and collector and stipulation of rules for accessing, tagging or copying data collected. Other directives include the collection limitation principle, the openness principle, and the individual participation principle (Directive 95/46/EC of the European Parliament and of the Council, 1995). The collection limitation principle specifies that the information should be collected fairly and not in hidden or manipulative ways, while the openness principle stipulates that the data subject should have knowledge of the collection of the data and have consented to it. Further, data subjects should have a right to know what of their personal information is stored and what it actually is; they also have the right to object to its retention and use. The individual participation principle allows data subjects to object to the use of their information. Certain categories of information should not be collected, such as race or ethnic origin, unless they are required to pursue the use of the data. Security and confidentiality of the personal information are also mandated, as is the notification of appropriate authorities as to the data mining and use.

3.1.3 OECD Fair Information Practice (FIP) Principles

The Organization for Economic Co-operation and Development (OECD) developed a set of guidelines in 1980 and revisited them in 2011 (Organization for Economic Co-operation and Development, 2011). The organization crafted eight principles that are similar to the other provisions listed here. The collection limitation principle speaks to how the data is collected, including whether knowledge and consent of the data subject was obtained. The data quality principle requires that personal data should be relevant, accurate, complete, and timely to the research purpose. The third principle, the purpose specification

principle, mandates that, at the time the data is gathered, the subject be told the purpose of current and subsequent research. The use limitation principle asserts that personal data not be disclosed or made available for purposes other than those specified in the purpose specification principle (with certain exceptions). The fifth principle, the security safeguards principle, mandates that data collected must be protected by reasonable measures against loss, destruction, unauthorized access, and so on. The openness principle requires that there be clarity and freedom of knowledge regarding practices and policies dealing with data. The seventh principle, the Individual Participation Principle, requires that individuals have rights to know if someone has data about them, what that data is in a reasonable time, with only reasonable costs, in a reasonable manner, and in a manner that is easily understandable. In the event that the data collector denies any of these claims, the data subject should be able to challenge the denial. Finally, the accountability principle requires that data miners are accountable for complying with the principles of the OECD policy.

3.1.4 The Code of Fair Information Practices

The US has adopted the Code of Fair Information Practices as a product of the passage of a myriad healthcare laws (Federal Trade Commission, n.d.; Electronic Privacy Information Center, n.d.). The changes in the healthcare laws that focus on the security of patient information is also applicable to the data mining of patient data and to data at large. The code has five principles. The first is that no personal data record-keeping systems should itself be kept secret. There should be notice that information is being collected and awareness on the data subjects' part that their data is indeed being collected, by whom, and for what purpose. Additionally, there must be a mechanism for the data subject to determine with whom the information may be shared. Further, this first principle dictates that one must solicit the information voluntarily; if it is not, the data subject must be made aware of the consequences of not sharing the information. The second principle is the idea that there must be a way for data subjects to find out what of their information is stored and how it is used: choice and consent. This provision sets out the opt-in or opt-out concept. Third, individuals should be able to access their personal information others hold and to correct inaccuracy or completeness. These abilities should be such that they allow timely, easy, and inexpensive avenues for the data subject to access and correct any mistaken or partial information.

The fourth principle relates to the data's integrity and security. The data should be accurate and securely maintained. Data collectors should take steps to assure that they have collected accurate information about the subject. Further, both technical and managerial steps should be taken to prevent loss, destruction, or unauthorized viewing of the data and to ensure the data is used in the way(s) agreed to. Finally, the fifth step involves enforcing provisions and measures and redress for those whose privacy has been violated. The code includes three remedies: self-regulation, private remedies, and government enforcement.

The review of all of these schemes to protect people's privacy seems to indicate a trend with regard to the chief characteristics of good codes of conduct for data mining. Fule and Roddick (2004) and Jakobsson, Juels, and Ratkiewicz (2008) summarize these characteristics: secure data sharing, confidentialization of publicly available materials, anonymization of private data, data subject control of access, notice and consent, and auditability. Pearson (2009) provides nine "key privacy requirements": notice, openness and transparency, choice, consent and control, scope/minimization, access and accuracy, security safeguards, (challenging) compliance, purposeful limitation of use with disclosure and retention policies, and accountability.

Finally, Jensen and Potts (2003, 2004) present a "practice of fair warning" for businesses using data mining with whom the consumer deals; it is based on three principles. First, the warning should be readily available to anyone affected by the data mining. Second, those affected should be given a mechanism for voicing concerns or questions about the data mining and their privacy rights. And, lastly, the warning should be one that those affected should be able, reasonably and in good faith, to understand what is being communicated. Table 2 summarizes all these major schemes and unifies the principles.

3.2 The Moral Perspective

The moral perspective is broader than the legal perspective and is what view is used in developing of code of ethics for online privacy policies. Cook and Cook (2003), van Wel and Royakkers (2004), Raiborn and Payne (1990), and Payne and Landry (2012), just to name a few, have all advocated for the use of the spirit of the law, rather than merely strict adherence to the letter of the law. The letter of the law is as strong, specific and binding as imperfect legal systems can make them, but there always seems to be

room for the unscrupulous to manipulate the law, in bad faith, to serve their own ends: in essence to violate, not the law itself, but the spirit of the law or the whole reason the law was passed in the first place. Rather than grant immunity from allegations of wrongdoing because of a poorly written law or laws that are, like issues dealing with privacy rights, inherently difficult or impossible to write comprehensively, clearly, etc., society can impose moral restrictions on behavior. Especially in business, moreover, consumers have the ultimate punishment for violations of the spirit of the law: they can withhold their patronage, a mighty tool to require compliance.

3.3 Culture and Moral Uncertainty

The importance of culture in discussing any issue with ethical overtones is certain. Culture includes shared values, beliefs and attitudes. It is unique to particular groups, a "collective programming" of the minds of the group members that differentiate it from other groups (Hofstede, 1980; Ma, 2010, p. 124). An individual's culture or cultures must be considered in developing and discussing online privacy policies. The possibility of a moral dilemma or moral ambiguity is also strongly influenced by culture. Values must also be defined before they can be used in constructing a model code of ethics for online privacy policies. Velasquez (1999) and Joyner, Payne and Raiborn (2002) have defined value as attributions of worth or that which is worthy or important in the decision making process. Values indicate socially or personally desirable elements.

Table 2. Online Privacy Policies: The Principles

ISTPA	EU directives	OECD	FTC	Common concepts
	Prior permission to store and access personal data	Collection limitation principle	Notice and awareness	Notice and awareness that information is being collected and used
	Notice and choice		Notice and awareness	
Security principle	Confidentiality and security principle	Security principle	Security	Sensitive information must be kept securely via technical and managerial means
Notice and choice principle	Information should be adequate, relevant, not excessive	Data quality principle Openness principle	Choice / consent	Consumers have a choice to provide or not to provide information and must consent to the sharing or use of the information Information should be accurate, relevant, complete and time-appropriate
Retention principle	Retained with regard to timeliness		Integrity	
Data integrity principle	Respect rights of subjects	Purpose specification principle	Integrity	Information should only be utilized for the purposes for which it was requested and / or approved
	Right of access	Individual participation principle	Access, participation	Access should be granted to the data subject to check for/correct mistakes, incompleteness, timeliness of data
Disclosure principle		Use limitation principle		Only data subject-approved sharing of data should be allowed

Cook and Cook (2003) define ethics as standards of conduct that cultures and organizations agree on; citing U.S. Supreme Court Justice Stewart, they state that ethics involves knowing the difference between what you have a right to do and what is right to do. Ethics springs from values that groups and individual members of a culture hold. Carroll (1979) and Freeman and Gilbert (1988) define ethics as an understanding of what is right and fair conduct or behavior. One can use the words ethics and morals as synonyms. Relative to online privacy policies, the kinds of information available and the ways information can be used provide a background set of knowledge with which to examine the moral perspective. Figure 2 represents the difficulty of knowing what information can be used and what should not be and gives some sense of the difficulty of knowing appropriate or inappropriate uses of that information.

3.4 Moral Codes of Conduct

We now turn to the morality literature to form the basis of our online privacy policy code of ethics. First, in seeking the values and ethics already present in extant codes of conduct for online privacy, we used two frameworks as the basis for the proposed code of ethics for payday lenders: Kant's categorical imperative (Kant, 1964) and Aristotelian virtues. Kant's categorical imperative includes three questions; if all of the answers are yes, a moral duty is imposed to act or not to act. First: is the action universally consistent; that is, would one chose to act such that all will be treated the same and such that the actor would submit to that treatment? For online privacy, the question could be phrased by the data collector: if this was my personal information, how would I want it collected, stored, and/or shared with others? Second: does the action respect individuals as inherently or innately valuable apart from any benefit that they might provide the actor? Again, for the data collector, the question might be whether collecting or using the data could somehow benefit the data subject rather than merely using the subject to satisfy some purpose only helpful to the data collector. Third: does the action acknowledge and respect the autonomy of all rational beings? The data collector should be aware that the consumer has the right to make a choice about providing data; its use, storage, and security; and how it is shared with other data miners. In the vernacular, we might state the categorical imperative as: one should do unto others as he would have others do unto him.

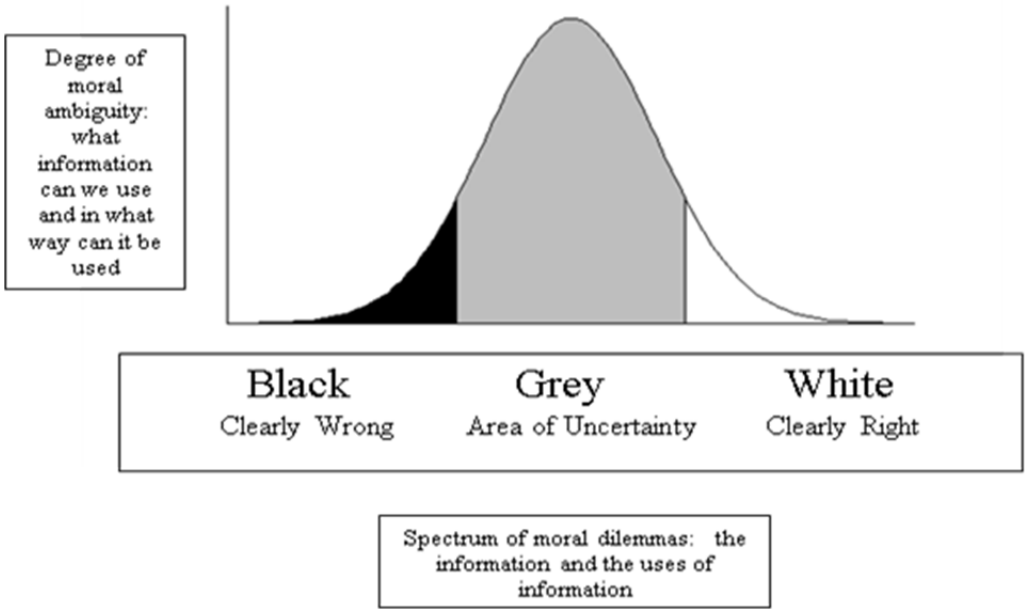

Figure 2. Proportionate Depiction of Areas of Moral Ambiguity

The Aristotelian virtues (Barnes, 1984; Bragues, 2006) also provide a construct on which to base our proposed code. There are six Aristotelian virtues of merit for online privacy policies: courage, self-control, generosity, magnificence, justice, and sociability. Courage is the ability to regulate fear, which should be present in reasonable measures when considering what would likely be seen by many as a questionable action, such as the sharing of personal information without the data subject's knowledge or consent. Self-control reflects attitudes towards pleasure and self-gratification. In a consumption-oriented society, both data miners and data subjects might succumb to excess: the data miner might inappropriately attempt to sell personal information about subjects, who may have given away too much personal information because they really wanted to complete the transaction, even knowing they were sharing too much.

In a similar vein, misusing the third virtue of generosity, which deals with the concept of attaining wealth, could be damaging to both miner and subject because the data miner might chose to take advantage of other people's vulnerabilities (i.e., the subject's need and want to engage in online transactions). In exceeding standards of conduct, data miners may sacrifice their reputation for meager returns (Bragues, 2006). Data subjects, on the other hand, can abuse the ideal of generosity by sharing too much information in an uninformed or careless way (Bragues, 2006). In either instance, there should be a median path that will suit the needs of each party: the lender to earn a reasonable return on investment and the borrower to have access to necessary funds. Magnificence is closely related to the concept of

generosity and implies expending large sums in the right way for a good reason rather than to be extravagant or boastful. Data miners could use the concept of magnificence to defend the amount of information sought or sold to a third party as being “reasonable”.

Sociability is the ideal that one should act pleasantly and professionally with others. Clearly, this attribute should permeate business dealings regardless of those dealings’ nature. This virtue is especially important in online privacy policies: data collectors should provide a venue for data subjects/consumers wherein providing personal information is less risky, safer, reasonable in depth and scope, and so on for the subject, who may be embarrassed about this private, “secret” information. Likewise, data subjects should be professional in assessing the need to share personal information and what kinds and amounts of personal information they should/want to share. Finally, although there are many levels to Aristotle’s discussion of justice, in its most simplistic form, justice reflects the idea of properly allocating goods. Aristotle believed that people should recognize the true value of exchanged items and that developing such information is part of the process of developing and adhering to a code of ethics (Bragues, 2006). Table 3 synthesizes philosophical attributes derived from Kant’s categorical imperative and the Aristolean virtues as they apply to online privacy.

Table 1. Philosophical and Principle Synthesis

Philosophical rule	Online privacy policy principles’ common concepts
Universally consistent actions	Sensitive information must be kept securely via technical and managerial means
Respect individuals as inherently valuable	Access should be granted to the data subject to check for/correct mistakes, incompleteness, timeliness of data
Respect autonomy of all rational beings	Consumers have a choice to provide or not to provide information and must consent to the sharing or use of the information
Courage	Information should be accurate, relevant, complete, and time-appropriate
Self-control	Data miners should be legally and socially accountable for failures
Generosity	Only data subject-approved sharing of data should be allowed
Magnificence	Information should only be utilized for the purposes for which it was requested and / or approved
Sociability	Notice and awareness that information is being collected and used
Justice	All of the above speak to the virtue of justice

4 Examining the Components for an Ethical Code of Conduct

To develop a code of conduct that is derived from both legal and ethical frameworks, the online privacy policy principles’ common concepts that Table 3 outlines will need to be examined through the lenses of the components required for secure data mining: people, processes, and technology. This examination addresses the conflict of business versus consumers and describes the elements essential for a code of conduct that can be used for multiple stakeholders. The main stakeholders of data mining are people. People in this context include both the consumers that supply their data and the data miners that handle that data. Both need to be educated about using that data beyond policies. Quite often, policies and procedures are implemented without an explanation of the reasons why a policy is needed. The focus should be on the reasons why the policy is needed and the functions it serves and not on the consequences and penalties for non-compliance. This education can be largely done through security awareness and training initiatives.

When reviewing the eight listed online privacy policy principles’ common concepts, security awareness and training need to be addressed for each of the stakeholders mentioned and tailored to fit that audience and function. Employees should be trained on how to handle and protect sensitive information through collection, processing, and storage addressing items 1, 2, 4, & 5 in Table 4. Agreements related to data mining, which include end user license agreements (EULA) and non-disclosure agreements (NDA), can

also educate and notify people on properly using data. EULAs provide a vehicle for consent and apply to items 3, 6, & 8 in Table 4 and NDAs provide the rules for usage and apply to items 5, 6, & 7. The problem is that these documents are often lengthy and are written in complicated language that is not easily read or comprehended. If the purpose of these documents is really for understanding, notification, and comprehension, they should be written in a format appropriate for that audience. Processes are the interactions between people and technology and include administrative controls such as policies and procedures as mentioned above. These processes need to be documented, tested, and enforced and apply to all eight online privacy policy principles' common concepts as noted in Table 4.

Technology should enable and provide the control mechanisms to protect data throughout all stages of collecting, processing, transmitting, and storing it. Strong access controls that enforce individual account authentication, authorization, and accounting must be employed. When examining the eight online privacy policy principles' common concepts, strong access controls are needed for items 1, 2, 4, 5, 6, and 7 (see Table 4). The system should include data validation techniques to ensure data is complete and in the right format during collection (item 2). Data encryption is needed to protect data at rest (stored on a PC, server, external hard drive, USB stick, backup tapes, etc.) and data in motion (transmitted across the network) and applies to items 1, 5, 6, and 7. However, even if strong access controls are employed along with protecting data at rest and in motion, there is still the possibility that the data miner may deploy the data to an unsecure system such as a personal email account or jump drive to work remotely. In doing so, items 1, 5, 6, and 7 are lost. To counter this loss of control, secure remote access should be deployed so that employees do not have to circumvent the policies and controls in place to do their job.

In reviewing the people, processes, and technology aspects of a code of conduct, Table 4 illustrates the need for integration among these items. While a code of conduct should not compromise security through disclosures (for example: we secure data at rest with symmetric encryption using AES with a 128 bit key), it should dictate and mandate that the organization will use security awareness, training, and agreements to educate people, develop mature processes, and employ technology that allow secure data mining.

Table 2. Online Privacy Policy Principles' Common Concepts And People, Processes, and Technology

	People			Processes			Technology		
Online privacy policy principles' common concepts	**Awareness**	**Training**	**Agreements**	**Documented**	**Tested**	**Enforced**	**Access controls**	**Data encryption**	**Secure access**
1. Sensitive information must be kept securely via technical and managerial means	●	●			●	●	●	●	●
2. Access should be granted to the data subject to check for/correct mistakes, incompleteness, timeliness of data	●	●			●	●	●		
3. Consumers have a choice to provide or not to provide information and must consent to the sharing or use of the information	●		●	●	●	●			
4. Information should be accurate, relevant, complete, and time-appropriate		●			●	●	●		
5. Data miners should be legally, socially accountable for failures		●	●	●	●	●	●	●	●
6. Only data subject-approved sharing of data should be allowed	●		●	●	●	●	●	●	●
7. Information should only be utilized for the purposes for which it was requested and / or approved	●		●	●	●	●	●	●	●
8. Notice and awareness that information is being collected and used	●		●	●	●	●			

5 Conclusions

Organizations have a responsibility to protect data that they collect, process, and store. This responsibility goes beyond due diligence and should focus on due care and taking actions as opposed to merely researching concerns. From the analyses and comparisons we make here, we conclude that a uniform code of ethical conduct for online privacy policies is not only in order but is also feasible from a managerial and ethical perspective. This code of conduct must incorporate the people, processes, and technical components to be truly useful. Future work could examine in two different research streams. The first is to examine existing online privacy statements to identify which tenets outlined here as common concepts are being deployed. The second is to refine our individual conclusions into a set of best practices and prescriptions that practitioners could adopt and tailor to meet their own specific customers and needs.

References

Barnes, J. (1984). *The complete works of Aristotle: The revised oxford translation*. Princeton: Princeton University Press.

Azmi, I. M. (2011). Bioinformatics and genetic privacy: The impact of the Personal Data Protection Act of 2010. *Computer Law and Security Review*, *27*, 394-401.

Baumer, D. L., Poindexter, J. C., & Earp, J. B. (2004). Meaningful and meaningless choices in cyberspace. *Journal of Internet Law*, *7*(11), 3-11.

Bennett, C. J. (2000). An international standard for privacy protection: Objections to the objections. In *Proceedings of the Tenth Conference on Computers, Freedom and Privacy: Challenging the Assumptions (*pp. 33-38). New York, NY: ACM.

Bragues, G. (2006). Seek the good life, not money: The Aristotelian approach to business ethics. *Journal of Business Ethics*, *67*(4), 341-357.

Carroll, A. B. (1979). A three-dimensional conceptual model of corporate performance. *Academy of Management Review*, *4*(4), 497-505.

Carusi, A., & Jirotka, M. (2009). From data archive to ethical labyrinth. *Qualitative Research*, *9*(3), 285-298.

Collingwood, L. (2012). Privacy, anonymity and liability: Will anonymous communicators have the last laugh? *Computer Law and Security Review*, *28*, 328-334.

Cook, J. S., & Cook, L. L. (2003). Social, ethical and legal issues of data mining. In J. Wang (Ed.), D*ata mining: Opportunities and challenges* (pp. 395-420). Hershey, PA: Idea Group Publishing.

Cranor, L. F. (2003). "I didn't buy it for myself" privacy and ecommerce personalization. In *Proceedings of the 2003 ACM Workshop on Privacy in the Electronic Society* (pp. 111-117). New York, NY: ACM.

Cranor, L. F. (2005). Giving notice: Why privacy policies and security breach notifications aren't enough. *IEEE Communications Magazine*, 18-19.

Danna, A., & Gandy, O. H., Jr. (2002). All that glitters is not gold: Digging beneath the surface of data mining. *Journal of Business Ethics*, *40*, 373-386.

De George, R. T. (2010). *Business ethics* (7th ed.). Upper Saddle River, NJ: Pearson.

Directive 95/46/EC of the European Parliament and of the Council. (1995). Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTMLElectronic

Privacy Information Center. (n.d.). *The code of fair information practices*. Retrieved from http://www.epic.org/privacy/consumer/code_fair_info.html

Federal Trade Commission (n.d.). *Fair information practice policies*. Retrieved from https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-report-congress/priv-23a.pdf

Freeman, R. E., & Gilbert, D. E. (1988). *Strategic management: A stakeholder approach*. Marshfield: Pitman Publishing.

Fule, P., & Roddick, J. F. (2004). Detecting privacy and ethical sensitivity in data mining results. In V. Estivill-Castro (Ed.), *Proceedings of the 27th Australasian Computer Science Conference* (pp. 1-8).

Hoffman, L. (2012). Data mining meets city hall. *Communications of the ACM*, *55*(6), 19-21.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills: Sage.

International Security, Trust, and Privacy Alliance. (2001). ISTPA framework project. Retrieved from http://emoglen.law.columbia.edu/LIS/archive/privacy-legis/ISTPA-FrameworkWhitePaper013101.pdf

Jackson, J. (2002). Data mining: A conceptual overview. *Communications of the Association for Information Systems, 8*, 267-296.

Jain, Y. K., Yadav, V. K., & Panday, G. S. (2011). An efficient association rule hiding algorithm for privacy preserving data mining. *International Journal on Computer Science and Engineering*, *3*(7), 2792-2798.

Jakobsson, M., Juels, A., & Ratkiewicz, J. (2008). Privacy-preserving history mining for Web browsers. In *Proceedings of the Web 2.0 Security and Privacy*.

Jensen, C., & Potts, C. (2003). *Privacy policies examined: Fair warning or fair game*? Retrieved from http://hdl.handle.net/1853/3215

Jensen, C., & Potts, C. (2004). Privacy policies as decision-making tools: An evaluation of online privacy notices. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 471-478). New York, NY: ACM.

Joyner, B., Payne, D., & Raiborn, C. (2002). Building values, business ethics, and corporate social responsibility into the developing organization. *Journal of Developmental Entrepreneurship*, *7*(1), 113-131.

Kant, I. (1964). *Groundwork of the metaphysics of morals* (H. J. Paton, Trans.). New York: Harper and Row.

Kelley, P. G. (2009). Designing a privacy label: Assisting consumer understanding of online privacy practices. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1-6). New York, NY: ACM.

Koh, S. C. L., Gunasekaran, A., Thomas, A., & Arunachalam, S. (2005). The application of knowledge management in call centres. *Journal of Knowledge Management*, *9*(4), 56-69.

Kucan, B. (2009). *Privacy and data protection in the European Union.* Retrieved from http://www.net-security.org/secworld.php?id=8129.

Ma, Z. (2010). The SINS in business negotiations: Explore the cross-cultural differences in business ethics between Canada and China. *Journal of Business Ethics*, *91*, 123-135.

Malhotra, N., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, *15*(4), 336-355.

Mehta, N., & Dang, S. (2011). Temporal sequential pattern in data mining tasks. *International Journal on Computer Science and Engineering*, *3*(7), 2674-2678.

Milne, G. R., & Gordon, M. E. (1993). Direct mail privacy-efficiency trade-offs within an implied social contract framework. *Journal of Public Policy Marketing*, *12*(2), 206-215.

Organization for Economic Co-operation and Development. (2011). *Thirty years after: The OECD guidelines for privacy.* Retrieved from http://www.oecd.org/dataoecd/63/56/49710223.pdf

Payne, D., & Trumbach, C. C. (2009). Data mining: Proprietary rights, people and proposals' business ethics: A European review, *18*(3), 241-252.

Pearson, S. (2009). Taking account of privacy when designing cloud computing services. In *Proceedings of the 2009 ICSE Workshop on Software Engineering* (pp. 44-52).

Ponemon Institute. (2006). Benchmark study of European and U.S. corporate privacy practices. Retrieved from http://jp.whitecase.com/files/Publication/1e7a69e0-49e9-478e-abc1-303e107c4dd7/Presentation/PublicationAttachment/4a78432a-bd1f-4363-ab82-32fab1729a1e/Benchmark_Study_Privacy_Practices_updated.pdf

Raiborn, C., & Payne, D. (1990). Corporate codes of conduct: A collective conscience and continuum. *Journal of Business Ethics*, *9*(11), 879-889.

Randeree, E. (2006). Knowledge management: Securing the future. *Journal of Knowledge Management*, *10*(4), 145-156.

Smith, H. J., Milberg, S., & Burke, S. (1996). Information privacy: Measuring individual's concerns about organizational practices. *MIS Quarterly, 20*(2), 167-196.

Tsai, J. Y., Egelman, S., Cranor, L., & Acquist, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, *22*(2), 254-268.

Van Wel, L., & Royakkers, L. (2004). Ethics issues in Web data mining. *Ethics and Information Technology, 6*(2), 129-140.

Velasquez, M. G. (1999). *Business ethics: Cases and concepts*. Upper Saddle River, NJ: Prentice Hall.

Wilder, C., & Soat, J. (2001). The ethics of data. *Information Week*, 36-48.

About the Authors

Dinah Payne, Professor of Management, has been licensed by the Louisiana Bar Association since 1986 and at the University of New Orleans since 1988. Her teaching and research interests are in the fields of business ethics, domestic and international law and management. She has participated in many international teaching experiences, having taught at the UNO-Innsbruck Summer School for many years and having participated in the Semester at Sea Program. Additionally, she has participated in a wide variety of international learning experiences, including attending seminars and presenting research in Costa Rica, Mexico, New Zealand, Italy and Belgium. She has been awarded a number of teaching, research and service awards, including the Seraphia Leyda University Fellow Award, the UNO College of Business Administration "Favorite Professor" Award and Professor of the Year Award and the Gordon "Nick" Mueller International Service Award. Her chief professional joy is working with, teaching and learning from wonderful students. She has been published in many journals, including the *Journal of Business Ethics, Communications of the ACM, the Labor Law Journal, the Journal of Corporate Accounting and Finance, the Journal of Developmental Entrepreneurship and Global Focus.*

Brett J. L. Landry is the Interim Dean at the Satish & Yasmin Gupta College of Business and Associate Professor of Cybersecurity at the University of Dallas. For more than twenty-five years, he has worked in information security in the public and private sectors. He has taught and consulted in the US, Europe, Asia, and South America and has published and presented numerous articles in the areas of Cybersecurity, IT Ethics, IT Management, Network Architecture, and Disaster Recovery. He also holds numerous industry security certifications such as Certified Information Systems Security Professional (CISSP), Certified Ethical Hacker (C|EH), Certified Information Systems Auditor (CISA) and Certified in Risk and Information Systems Control (CRISC).

Matthew D. Dean is an Associate Professor in the School of Business at the University of Southern Maine. He teaches in the fields of management science, operations management, and statistical data analysis. The crux of his research involves developing modeling tools and methodologies for addressing complex management decision-making challenges. He has published in numerous journals, including *Operations Research, Decision Support Systems, and Communications of the ACM.*