6. CORE PRIVACY
A Problem for Predictive Data Mining

JASON MILLAR*

I. INTRODUCTION

Data mining technologies are pervasive in our society.[1] They are designed to capture, aggregate, and analyze our digital footprints, such as purchases, Internet search strings, blogs, and travel patterns in an attempt to profile individuals for a variety of applications. Knowledge of the full scope of data mining often leaves people feeling as though it is intuitively wrong. On the one hand, this feeling is written off just as an unreflective expression of unease toward the technology, the potential harms of data mining (assuming any can be identified) outweighed by its potential benefits. On the other hand, the feeling is often explained by an appeal to more academic conceptions of privacy. Built primarily on surveillance technologies, data mining is implicated as a potential violation of individuals' privacy. People's unease with data mining might best be understood as the expression of some intuitive understanding that their privacy is at stake.

Nissenbaum and Tavani address the privacy threats associated with data mining, and analyze them based largely on contextual accounts of privacy.[2]

[*] The author would like to thank Jacquie Burkell, Jeremy Clark, Carole Lucock, David Matheson, Val Steeves, and Ian Kerr for valuable insight and feedback regarding this article. Ian Kerr, Canada Research Chair in Ethics, Law and Technology at the University of Ottawa, deserves special thanks for providing countless students an opportunity to cut their academic teeth in distinguished company.

[1] Andy Clark, *Natural-born Cyborgs: Minds, Technologies, and the Future of Human Intelligence* (New York: Oxford University Press, 2003); David Lyon, "Facing the future: Seeking ethics for everyday surveillance," *Ethics and Information Technology* 3, no. 3 (2001): 171–181; Helen Nissenbaum, "Protecting Privacy in an Information Age: The Problem of Privacy in Public," *Law and Philosophy: An International Journal for Jurisprudence and Legal Philosophy* 17, no. 5–6 (1998): 559–596.

[2] See Nissenbaum "Protecting Privacy in an Information Age: The Problem of Privacy in Public," (n. 1).; H.T. Tavani, "Informational Privacy, Data Mining, and the Internet,"

But a privacy analysis of predictive data mining underpinned by contextual arguments is unable to fully analyze the unique privacy threats posed by the technology. The reason is that contextual accounts of privacy are open to two objections that predictive data miners can exploit in defense of their technology. As a result contextual accounts of privacy suffer in trying to describe what is problematic with predictive data mining where profiling individuals is the goal.

In order to better understand what *feels* wrong about data mining to many people, I will focus on some unique characteristics of predictive data mining. In particular, I will argue that current technology allows datasets to be analyzed with a level of sophistication that results in the emergence of new types of data. Fully analyzing the privacy implications of the new data requires first, an evaluation of what types of data count as *core private information* and second, an evaluation of what constitutes a *core privacy violation*.[3] By viewing data mining in this new light, I hope to explain, and overcome, one of the theoretical barriers to a full privacy analysis of predictive data mining, while providing a methodology for use in assessing, on a case-by-case basis, whether or not particular instances of predictive data mining constitute a privacy violation.

II. PREDICTIVE DATA MINING AND PROFILING THE INDIVIDUAL

It is impressive to think of the number of digital footprints each of us will have left behind during the course of our lives. Every swipe of an electronically readable document (e.g., credit cards, ATM cards, student cards, library cards, passports, etc.), every search string entered into a Web browser, every blog entry and email, any Electronic Product Code (EPC) identified by a reader, all phone calls and connections between an individual's digital hardware and a network, and every other interaction between an individual or her possessions and a computer of some sorts, is logged and time stamped in an electronic database for possible future reference.[4]

Ethics and Information Technology 1 (1999): 37–145.; and H. T. Tavani, "KDD, Data Mining, and the Challenge for Normative Privacy," *Ethics and Information Technology* 1 (1999): 265–273.

[3] Within the Canadian legal system the notion of "a biographical core of personal information" has been used to determine whether or not an individual has a reasonable expectation of privacy with respect to certain information (see *R. v Plant*, [1993] 3 S.C.R. 281.). Though there may be overlaps between what counts as core private information or as a core privacy violation (i.e. those concepts I intend to develop) and the legal notion found in Canadian jurisprudence, they are not intentional. My intentions are philosophical, directed at understanding how to think of privacy in light of particular data mining technologies. If there are legal implications buried in my account, I leave it to those better versed in the law to decide what they may be, and elucidate (or import) them accordingly.

[4] EPCs are currently in the advanced stages of development by EPCglobal, Inc (see http://www.epcglobalinc.org/home). They will be deployed in the form of Radio Frequency

A great majority of those footprints are either explicitly associated with identifiable individuals or were designed to make the process of identification relatively painless. That means that the databases collecting the footprints can contain identifying information about the individuals alongside the footprint data. Exceptions to this are health care information, which is intentionally "anonymized," and some personal data sold to third parties that, depending on the legal jurisdiction, must be "anonymized" prior to release.[5] In a given year, a conservative estimate of twenty digital transactions a day means that more than 7,000 transactions become associated with a particular individual—upwards of a half million in a lifetime. Capturing and storing this information is becoming easier and cheaper—the number and detail of our daily transactions is increasing. Actual numbers need not be calculated in order to get the sense that the resulting datasets represent a source of great potential value to anyone interested in analyzing, or *mining*, them.

The field of data mining is burgeoning. There are no fewer than thirty major technical conferences held internationally each year.[6] According to the Institute of Electrical and Electronics Engineers (IEEE) International Conference on Data Mining official Web site, the number of paper submissions for the conference's proceedings more than doubled to 776 in its first six years of being held.[7] Various topics are discussed at these technical conferences including new algorithms for use in solving specific problems and Knowledge Discovery in Databases (KDD), which is defined as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data.[8] Despite the growing size of the literature and increasing sub-specialization, data mining can generally be understood as the search for patterns within a given dataset.

Identifier Tags capable of uniquely identifying each individual product (as opposed to a box or flat containing several of the same product) that rolls off an assembly line, by means of a 96-bit number.

[5] This should not give the impression that the so-called "anonymous" data cannot be used to re-identify individuals. B. Malin, L. Sweeney, and E. Newton, "Trail Re-Identification: Learning Who You Are From Where You Have Been," in *Carnegie Mellon University, School of Computer Science, Data Privacy Laboratory Technical Report, LIDAP-WP12* (Pittsburgh: Carnegie Mellon University, Laboratory For International Data Privacy, February 2003), point to the ease with which most such data can be re-associated with the individuals they belong to.

[6] IEEE International Conference on Data Mining, http://www.kmining.com/info_conferences.html (accessed March 5, 2008).

[7] Last year of data was the 2006 conference at the time of writing. IEEE International Conference on Data Mining, http://www.cs.uvm.edu/~icdm/ (accessed June 17, 2008).

[8] W. J. Frawley, G. Piatetsky-Shapiro, and C. Matheus, "Knowledge Discovery In Databases: An Overview," in *Knowledge Discovery In Databases*, eds. G. Piatetsky-Shapiro and W. J. Frawley (Cambridge, MA: AAAI Press/MIT Press, 1991): 1–30.

There are two broad categories of data mining tasks: *descriptive data mining*, where the goal is to describe the general properties of the existing data, and *predictive data mining*, where the goal is to predict based on inference from the original data.[9] Descriptive tasks might include comparing the features of two datasets as in determining that they are identical or similar to some degree, or can involve simple characterizations like "dataset *x* includes fifty-two instances of doughnut purchase activity." Predictive tasks, on the other hand, include activities like *profiling*. Inferences can be drawn from the regularities and patterns found in datasets, which allow certain predictions to be made. For example, data from hospital emergency room visits is mined such that potential outbreaks of diseases can be effectively predicted during the early stages of the outbreak. On a more personal level, individuals' Web browsing patterns can be predictively mined in order to glean the types of information that are of particular interest to those individuals. Based on the resulting dataset, predictions can be made about them including the types of music that they may like to purchase, political or religious affiliations they may have, or illnesses from which they may suffer.

A diverse group of individuals currently applies predictive data mining techniques during the course of its business. Law enforcement and intelligence agencies have long relied on data mining techniques to develop detailed profiles of individuals (political leaders, criminals, etc.) in an attempt to better predict the behavior patterns of those individuals based on past behavior patterns. They, and militaries, are increasingly making use of data mining techniques to identify potential threats. Credit card companies use data mining techniques to better detect fraudulent activities. Cross-referencing each transaction request with models of regular versus irregular activity, and flagging potential irregular activity have resulted in fewer losses due to fraudulent activity. Security companies use data mining to monitor and analyze the information captured on video cameras and other monitoring devices (swipe cards, etc.) in order to identify potential security threats or breaches. Businesses use data mining to profile customers and understand their potential purchasing preferences. Web sites use data mining techniques to better serve up information, such as search results and advertising, to those browsing their sites.

These descriptions of data mining underscore the important role that surveillance technologies play in the overall data mining scheme. Surveillance typically involves the targeted collection of seemingly disparate bits of information, via the application of various technologies, for potential use in some sort of post hoc analysis. Cameras collect data for analysis after some interesting activity has taken place; swipe card readers log the comings and goings of people in some

[9] O. R. Zaiane, J. Lia, and R. Hayward, "Mission-Based Navigational Behaviour Modeling for Web Recommender Systems," in *Advances in Web Mining and Web Usage Analysis: Lecture Notes in Artificial Intelligence*, eds. B. Mobasher, O. Nasraoui, B. Liu, and B. Massand, (Springer Verlag, LNAI 3932, 2006): 37–55.

secure area; phone logs are recorded for use in billing or potentially for use in criminal investigations. Surveillance technologies can be seen as the backbone of data mining, providing the seemingly disparate pieces of information that populate datasets. Those datasets are then fed into the various data mining algorithms and the search for patterns begins.

III. ASSESSING THE PRIVACY IMPLICATIONS OF DATA MINING

Data mining, both descriptive and predictive, carries privacy implications. An individual's privacy, in the most general terms, turns on the status—known versus unknown, public versus private—of particular kinds of information about an individual. When a certain piece of information deemed private by some theory, say theory *A*, becomes known about an individual, the individual is said to have suffered a privacy loss according to theory *A*. Data mining technologies provide a means of discovering information about individuals. So determining whether or not the process or results of data mining are privacy invasive is the crux of the problem at hand.

Before outlining some assessments of the privacy implications of data mining, a word about the public-private dichotomy of personal information is in order. The particular version of it that I will offer here conflicts with some theories of privacy. However, I mean only to expound what may be a typical version of the dichotomy for clarity; my overall argument does not depend on the correctness of this account, only on the fact that the public-private dichotomy of personal information is commonly the focus of privacy theories.

Given the above generalization of privacy theories we can say that bits of information about an individual may be considered public or private. An example of a public piece of information could be that person X was walking down Main Street last night; person X was in a public space and had no expectation of privacy with respect to his whereabouts. If another individual came to know that person X was on Main Street last night then person X would not suffer a privacy loss with respect to that piece of information. An example of a private piece of information could be that person X is secretly in love with person Y and expresses this love in the form of letters hidden in X's closet. If another person were to stumble across the cache of love letters and proceed to read them, person X would have suffered a loss of privacy with respect to the information about her feelings toward Y. One of the main problems that a theory of privacy needs to address is expounding what qualifies a particular kind of information as public or private.[10] Generally speaking, public information is considered to be

[10] See Judith Jarvis Thomson, "The Right to Privacy," *Philosophy and Public Affairs* 4 (1975): 295–314, and T. Scanlon, "Thomson on Privacy," *Philosophy and Public Affairs* 4 (1975): 315–322, for good examples of this approach to theorizing about privacy.

public full stop, and private information is considered to be private full stop, in other words the status of information about an individual is not vague.

How do we assess the privacy implications of data mining? Nissenbaum and Tavani have focused largely on problems related to the public versus private status of the data that is included in the original dataset used for data mining, and construct a contextual account of why data mining can violate privacy norms. In this section, I will argue that this has the effect of preventing them from fully assessing the privacy implications of predictive data mining, especially with regard to the resulting data, or knowledge, that is unique to predictive data mining.

A brief sketch of Nissenbaum's and Tavani's arguments is in order to properly assess their approach to the problem. Both of their arguments adopt a contextual conception of privacy; they focus on the importance of an individual's control over the flow of personal information because that information can have different privacy implications from one social context to another. Rachels's account of why privacy matters points to the root of their theories. He argues that privacy matters because it is the mechanism we use to maintain the integrity of our relationships with one another.[11] He says,

> Because our ability to control who has access to us, and who knows what about us, allows us to maintain the variety of relationships with other people that we want to have, it is, I think, one of the most important reasons why we value privacy.[12]

The claim that a contextual control over information is essential to privacy is echoed in a roughly equivalent form in each of Nissenbaum's and Tavani's arguments, which are explicit attempts to understand the privacy implications of data mining. Nissenbaum rejects the traditional public-private dichotomy of information (described above), arguing instead that certain bits of information relating to a person can be public in one context and private in another.[13] For example, knowledge that person X was walking down Main Street last night could be public with respect to a certain set of X's friends, but deeply private with respect to X's co-workers if, for example, Main Street happened to be in the heart of the local gay district and X was aware that his co-workers would interpret his mere presence there as cause for discrimination. According to Nissenbaum this dual nature of information is problematic where public surveillance and data mining are concerned because of the inherent potential for data to be taken out of context. Data mining is of particular concern because it frequently involves the

[11] James Rachels, "Why Privacy is Important," *Philosophy and Public Affairs* 4 (1975): 323–333.

[12] Rachels, "Why Privacy is Important," (n. 11).

[13] Nissenbaum, "Protecting Privacy in an Information Age: The Problem of Privacy in Public," (n. 2).

buying, selling, and trading of so-called public data for use in new contexts—contexts in which the data was not explicitly divulged and in which that data is private.

Tavani offers reasons for considering data mining a new threat to privacy, above and beyond traditional data retrieval techniques.[14] Though he appears to be analyzing the unique features of predictive data mining tasks (i.e., the discovery of unknown, nontrivial information), in assessing their privacy implications he seems to emphasize the contextual implications similar to those raised by Nissenbaum. Consider the privacy implications he draws from a scenario involving Lee, an individual who has divulged information to his bank while applying for a car loan. Lee's information is mined by successive algorithms, which in turn determine that Lee is part of a high credit-risk group of executives that is likely to declare bankruptcy within five years, despite current gainful employment. Tavani notes the following:

> Why does the mining of data about Lee by the bank raise concerns for privacy? While Lee voluntarily gave the bank information about annual salary, previous loans involving vacations, and the type of automobile he intended to purchase, he gave each piece of information for a specific purpose and use. Individually, each piece of information was appropriately given in order that the bank could make a meaningful determination about Lee's request for an automobile loan. However, it is by no means clear that Lee authorized the bank to use disparate pieces of that information for more general data mining analyses that would reveal patterns involving Lee that neither he nor the bank could have anticipated at the outset.[15]

Although Tavani focuses specifically on the particular strengths of data mining to uncover unknown patterns in a dataset, the example he uses suggests that the discovered knowledge is problematic from a privacy perspective because it shifts the context within which the data is considered. The implication seems to be that Lee would consider the data public with respect to his application for a car loan, but private with respect to a determination of his future credit risk.[16] Even though the data resulting from the predictive data mining is considered problematic, the privacy implications remain in relation to the dual public-private nature of the data in the original dataset.

[14] Note that those techniques are essentially traditional methods of surveillance in which the main activity is the collection and aggregation of physically observable data for future analysis; Tavani, "Informational Privacy, Data Mining, and the Internet," (n. 2).

[15] Tavani, "Informational Privacy, Data Mining, and the Internet," 141 (n. 2).

[16] There is an additional implication that the resulting risk profile is also problematic because it could be false. This is a common argument against data mining in general, and it, too, fails to allow for a full analysis of predictive data mining, which I will address shortly.

So Nissenbaum and Tavani each describe the privacy implications of data mining with a focus on analyzing the contextual aspects (i.e., the public-private status) of the original dataset. The data itself is described as moving from a broadly public kind, as in the information divulged by X during his walk down Main Street, or from a narrowly public kind, as in the information divulged by Lee during his car loan application, to a private kind owing to some shift in the context of analysis.[17]

If our goal is to fully assess the privacy implications of predictive data mining, then I think we face a theoretical difficulty if we limit our arguments to the contextual nature of information. There are at least two problems with this approach. First, both Nissenbaum and Tavani mention how contextual arguments seem to be open to a normative "knock-down" objection, namely that profiling is acceptable because the individuals whose privacy is supposedly at stake publicly divulged all of the information in the original dataset.[18] This is a formidable objection that Nissenbaum and Tavani attempt to address in each of their arguments; if individuals have publicly divulged information that is subsequently mined, how can they make a post hoc claim to privacy with respect to information *gleaned only from that dataset*? As I have pointed out, Nissenbaum and Tavani both offer useful arguments how this can be possible. In addition, Nissenbaum claims that contextual norms are violated during profiling, which places the public at risk of manipulation and signals a loss of control of their personal information. These two factors, she suggests, are part of the reason that data mining (and public surveillance) is ill received by the public. However, even in the wake of their analyses, the objection remains formidable against their (and any other) contextual account of privacy.

A second problem arises with a contextual analysis of data mining. It is the possibility of a scenario in which data mining is performed that raises no contextual privacy concerns of the kind that may result from a shift in context of analysis, and in which the results turn out to be accurate. One could imagine that even in that scenario knowledge resulting from predictive data mining could be privacy invasive, a fact that would depend only on the nature of the resulting knowledge, though not on its being false. A thorough privacy analysis of predictive data mining must account for these scenarios in addition to those covered adequately by contextual analyses.

Though they certainly provide important insight into privacy problems associated with the flow of information due to data mining, Nissenbaum's and Tavani's arguments each seem to fall short of allowing a full assessment of the unique aspects of predictive data mining associated with the discovery of

[17] "Broadly" and "narrowly" referring to the relative expectation of privacy that X has with respect to the divulged information.

[18] Nissenbaum, "Protecting Privacy in an Information Age: The Problem of Privacy in Public," 587 (n. 1).

new knowledge. Focusing on the nature of the knowledge discovered through the process of predictive data mining, rather than on the contextual (public-private) status of the data in the original dataset, might add to our understanding of privacy in a way that allows us to further explain the bad taste that data mining leaves in so many of our mouths.[19] To this end, the unique characteristics of predictive data mining technologies, as well as the unique nature of the data that is produced by it, must be directly addressed.

IV. THE EMERGENT DATA OF PREDICTIVE DATA MINING

Focusing on the contextual nature of data in a privacy analysis of predictive data mining only tells us part of the story. What can be said of the kinds of data that are the results of complex predictive analyses where the goals include such things as psychological profiling?

Predictive data mining tasks, including those related to the psychological profiling of individuals, rely on KDD for their resulting data. Recall that KDD is characterized as the nontrivial extraction of unknown information from datasets. In the case of psychologically profiling individuals, the types of resulting knowledge that are of most interest are those offering the best predictions about the individuals' underlying psychological properties (i.e. the individuals' beliefs, desires, or intentions). This is because those kinds of psychological properties are typically seen as the causal roots of action production. Thus individuals' underlying beliefs, desires, or intentions combine to motivate their purchases, crimes, or mouse clicks. Discovering the kinds of knowledge—from seemingly disparate pieces of information in a dataset—that most accurately reflect an individual's psychological properties is therefore the holy grail of psychological profiling via predictive data mining.

Is it reasonable to expect that a computer algorithm could somehow provide accurate representations of an individual's psychological properties? It is beyond our current theoretical landscape to answer in the affirmative as we lack the scientific and philosophical knowledge of the mind that would allow us to clearly delineate beliefs from desires from intentions, or to give an accurate account of what those properties consist of from a physical or psychological perspective. But this does not preclude us from articulating the privacy implications of algorithms that attempt to discover knowledge about an individual's beliefs, desires, or intentions. Describing the precise relationship between the knowledge resulting from predictive data mining and our mental properties may not be

[19] Nissenbaum, "Protecting Privacy in an Information Age: The Problem of Privacy in Public," (n. 1), argues that proof of the bad taste can be found in a 1990 poll, which showed that consumers felt that they were being asked to provide "excessive personal information."

necessary in order to begin articulating a theory of privacy that accounts for predictive data mining activities. Any talk of data representing psychological properties from here on in should therefore not be interpreted as a statement about an identity relationship between the two. By considering specific examples where psychological profiling is applied, I will argue that we can identify cases where predictive data mining is practically successful in its attempt to represent an individual's beliefs, desires, or intentions. I will also provide a means for gauging the success of a psychological profiling algorithm that is meant to address the underlying unease that may be felt concerning predictive data mining.

Consider an example in which predictive data mining is applied to Web browsing activity to determine what particular kind of information an individual is searching for. Predictive data mining here is applied in an attempt to satisfy her informational *intentions or desires*. A dataset composed of the content of Web pages she has visited, her Web access history and the connectivity between resources in a Web site can be used to predict the particular kind of information she is currently looking for.[20] Resulting data, namely the predictions about what she might currently be looking for, can be typologically distinguished from the data contained in the original dataset—the new data is meant to reflect her intentions or desires, whereas the original dataset contained seemingly disparate pieces of information divulged by virtue of her Web browsing activity. The dataset used in the prediction could theoretically be composed not only of an individual's current sessional data, but also of her entire Web browsing history, as well as knowledge of her blogging activity and any other piece of information about her that is made available to the algorithm. An example of this kind of predictive data mining project is Google's Web History (formerly Google Personal Search), designed to deliver personalized search results based on search string input and the individual's entire Web search and click history.[21] In addition to recommending information, data mining could be performed on such a dataset to try to predict a host of other information, the individual's religious beliefs and political beliefs among others.

This account demonstrates how psychological profiling differs from descriptive data mining in an important respect: the resultant dataset emerging from this kind of predictive data mining contains new data items that are not mere aggregations or simple characterizations of the original dataset. Those new data

[20] This particular kind of dataset is described in O. R. Zaiane, J. Lia, and R. Hayward, "Mission-Based Navigational Behaviour Modeling for Web Recommender Systems," (n. 9), as a useful dataset for use in Web recommender systems (i.e., those systems that recommend particular options to users based on previous Web activity).

[21] See http://www.google.com/psearch for Google's own description of the product. Andy Clark, *Natural-born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*, (n. 1) also describes in detail the kind of qualitative difference between this kind of technology and the more traditional search-string-based information recommender system.

items are better characterized as non-trivial representations of an individual's beliefs, intentions, or desires. Hence, the resultant dataset includes emergent data that must be considered in its own right if we are to properly characterize predictive data mining and the privacy implications of it.

Understanding the nature of the emergent data contained in the resulting dataset can be aided by a trivial case of predictive data mining, one that may be carried out by any attentive individual. Imagine your coworker, Jon, shows up for work every day eating a chocolate doughnut. Based on the dataset including only the information, 1) that Jon walks by every day with the same thing in his hand, and 2) that the thing is always a chocolate doughnut, one could descriptively conclude that "Jon eats a chocolate doughnut every work day." However, being human and knowing that Jon is also human, one may also use the data to draw the predictive conclusion, "Jon likes chocolate doughnuts," thus attributing a desire, or preference, to Jon that is not a mere description of the original dataset since the original dataset does not contain any data about Jon's preferences or desires. A statement about Jon's desires is qualitatively different from the data in the original dataset, and that qualitative property is one that emerges by virtue of a trivial yet complex prediction about Jon's psychology.

Recall that my concern is to address the scenario where contextual accounts of privacy fail to provide an adequate means of dealing with the resulting knowledge discovered by a predictive data mining algorithm. It would help to have a methodology capable of dealing with the scenario in which a predictive data mining algorithm produces highly accurate results about an individual. Some privacy arguments claim that it is the falsity or incompleteness of the predictions that prove problematic, and they are easily refuted by those claiming to have accurate results.[22] So an account of how we might gauge the success of predictive data mining algorithms is in order. In the Web browsing case, partial success of the profiling might be gauged by the individual's willingness to click on a link to the recommended information. Complete success, however, could be gauged by her complete satisfaction with that recommendation. In other words, if an individual has a particular informational intention that is completely satisfied by the data resulting from the predictive algorithm, we could say that the profile correctly represented her intentions in the form of a recommendation.

It could be objected that the emergent data are superficially coincident with psychological properties at best. For example, a Web recommender application might mine an individual's music collection and browsing habits (e.g., on an online music store) to pick out similarities between artists that an individual likes in order to recommend artists the person *might* like. But the assertion that those recommendations are anything like a person's beliefs, intentions, or

[22] David Lyon, "Facing the future: Seeking ethics for everyday surveillance," (n. 1), seems to use this argument in the way that he claims an individual can become alienated by the "faceless" profiles resulting from data mining.

desires remains unfounded. If it turns out that the individual does like the recommendations, then the recommendations are a statement about the predictability of his listening habits, or some other mere statistical fact about musical purchase trends in general. A person's beliefs, desires, and intentions need not be inferred in order to predict, or suggest with high rate of success, his purchases. Therefore, the psychological assertion is spurious.

What counts for an analysis of the success of predictive data mining is the *psychological resemblance* between the prediction and the individual's underlying psychological properties. If the emergent data of predictive data mining psychologically resembles an individual's beliefs, intentions, or desires, then we can gauge the data mining a success and put the objection to rest. Psychological resemblance between an output of a psychological profiling algorithm and an individual's beliefs, intentions, and desires, is something we can assess empirically. Our means of accomplishing this borrows from a well-known example in artificial intelligence—the Turing test.

Turing proposed his test as a means of recognizing intelligence in a machine.[23] In order to tell if a machine is intelligent, he suggested, an interrogator poses a series of questions to one of two potential respondents: a human or a machine. The interrogator is unaware which he is questioning; the identity is somehow hidden. His task is to determine if he is interrogating the machine or the human. If he is unable to correctly identify the respondent (with some statistically significant reliability), then the machine is said to be as intelligent as a human being. In the Turing test, the inability to distinguish between artificial intelligence and human intelligence counts toward asserting the machine's intelligence even though it is *not a statement about the identity between the kind of mental activity common to human intelligence and that activity occurring in the machine.*[24] A similar distinction can be exploited to provide a means of gauging the success of a predictive data mining task whose goal is to psychologically profile an individual by means of producing data resembling his beliefs, desires, or intentions.

Turing's test can be modified to test for psychological resemblance between the emergent data of predictive data mining and psychological properties, thus gauging the success of a predictive data-mining task. Turing's approach is a good candidate for our purposes particularly because the interrogator, in our case, has first person access to the comparison class against which the emergent data of predictive data mining is tested, namely his own beliefs, desires, and intentions. We can define the test as an attempt to determine whether the emergent data is considered *synonymous* to his corresponding psychological properties.

Our interrogator, in this test, would approach a prediction from data mining in a self-reflective inquisitive mode. Take a predictive data mining algorithm

[23] Alan M. Turing. "Computing machinery and intelligence," *Mind.* 59, no. 236 (1950): 433–460.

[24] Daniel Dennett, "The Age of Intelligent Machines: Can Machines Think?" KurzweilAI.net, http://www.kurzweilai.net/articles/art0099.html.

designed to determine political beliefs as an example of how the test would run in determining the psychological resemblance of a data mining prediction to a belief. Our interrogator is submitted to whatever surveillance method is required to populate the dataset used by the algorithm, say by turning over his entire Web browsing and search history. The interrogator is then asked to describe, in as much detail as possible, whatever political beliefs he possesses in accordance with whatever beliefs the data mining algorithm will target as predictions. The algorithm then predicts his political beliefs or more particular political beliefs, maybe in relation to one or several particular hot button issues. If the prediction matches the interrogator's own description of the belief (i.e., if the interrogator is satisfied to a sufficient degree that the prediction is what he believes) then it qualifies as synonymous to his actual belief.

I will refer to this as the *synonymy test*: if an interrogator is unable to distinguish the emergent data from a self-generated description of the target psychological property of the prediction (e.g., the particular belief, intention, or desire) to a sufficient degree, then the data and psychological property qualify as synonymous. Cases satisfying the synonymy test would be cases where the emergent data *psychologically resemble* the interrogator's psychological properties, the upshot being that they are cases of successful predictive data mining.

Why is this kind of synonymy sufficient for gauging the psychological resemblance between emergent data and a psychological property? Let us suppose that an interrogator produces a self-generated description of whatever the target of the data mining algorithm is. Let us suppose, then, that she is reasonably convinced that some particular set of emergent data is such that her best self-generated description of her beliefs, desires, or intentions represented by those data is synonymous with them. For the interrogator to then claim that they do not resemble one another would seem to be a case of incredulous question begging. At the very least, she would need to provide some description of how they do not resemble one another to support her rejection of their resemblance, especially since she has already agreed that they are synonymous according to the test. Of course, if such a description were then produced it would presumably qualify as self-generated, and would provide descriptive information individuating it from the emergent property, meaning that she was initially mistaken about the synonymy.

This approach to gauging success could be considered problematic because of a well-documented psychological phenomenon known as the Forer Effect (or Barnum Effect, or the Fallacy of Personal Validation). According to the Forer Effect, people tend to consider vague statements about personality, such as those found in horoscopes, to be highly accurate and unique to them, even though they could apply to almost anyone.[25] Given that the synonymy test relies on the individual's own evaluation of whether or not the prediction relates to her to a satisfactory degree, the systematic nature of the Forer Effect could render every

[25] D. H. Dickson and I. W. Kelly, "The 'Barnum Effect' in Personality Assessment: A Review of the Literature." *Psychological Reports*, 57 (1985): 367–382.

prediction highly synonymous, thus rendering the test useless. In addition, it could have a similar force to the objection raised above, in that the Forer Effect would limit our statements about synonymy and psychological resemblance to statements about mere statistical facts, thus undermining the stronger psychological claim.

There is a problem with this objection. The Forer Effect describes how an individual incorrectly believes that a vague descriptor (like a typical description contained in a horoscope) uniquely applies to him.[26] As such, horoscope-like descriptions might actually pass the synonymy test, as would a host of other trivial descriptions—that X likes to read a good book now and again, or that X believes murder is wrong—but this does not pose a significant problem to the synonymy test in the context of predictive data mining. Practically speaking, it would be unusual for anyone to question the privacy implications of a horoscope. Regardless, a predictive data-mining algorithm that provided vague, horoscope-like descriptors as output would likely be considered a failure, especially since the goal of KDD is to discover *nontrivial* knowledge. Nontrivial information—an individual's voting intentions, sexual history, or health profile—is not attributable to broad swaths of the population, making it unlikely that a random individual would incorrectly adopt it in self-description. Indeed, the less trivial the descriptor, the less likely it seems that the Forer Effect would obtain and invalidate the synonymy test.

By focusing on the unique results of predictive data mining designed for the discovery of nontrivial knowledge for psychological profiling, we are faced with assessing the privacy implications of a new kind of data. We can understand the relationship between emergent data and the target individual's psychological properties in terms of their psychological resemblance. Psychological resemblance can be used as the relevant test in assessing the success of a predictive data mining task aimed at psychological profiling. Psychological resemblance finds its natural ethical implications under the rubric of *core privacy*.

V. CORE PRIVACY

> There is much more going on inside us that we are willing to express, and civilization would be impossible if we could all read each other's minds.[27]

If it is possible to identify a kind of information that is essentially private, the emergent data of predictive data mining is a good place to start looking. Despite our intuitions that we often have something like a right to privacy regarding such information, the public nature of the original dataset provides for a formidable

[26] Forer reportedly constructed his personality descriptors from horoscopes.

[27] Thomas Nagel, "Concealment and Exposure," *Philosophy and Public Affairs,* 27, no. 1 (1998): 3–30, 4.

objection to privacy claims—how could someone who rightfully obtained the information subsequently use it and pose a threat to privacy in doing so? In addition, accuracy in the resulting data further complicates objections to data mining based on the emergence of false positives. Contextual accounts of privacy go a long way in describing some of the privacy issues related to data mining, but they fail to fully address how emergent data that is both obtained and used in the same context, and is highly accurate, could threaten privacy. Providing a test for assessing the psychological resemblance between emergent data and an individual's target psychological properties allows us to gauge the success of a predictive data mining algorithm. In this section, I will outline the privacy implications of emergent data. In doing so, I will argue that we can construct an understanding of *core privacy*, and that certain information counts as *core private information*. By focusing on how emergent data potentially violates core privacy, I hope to provide an argument for understanding why predictive data mining is intuitively, and practically, problematic.

Nagel's claim suggests a point of departure for assessing the privacy implications of emergent data. Each of us entertains a great number of thoughts (psychological properties) that we choose not to divulge for some reason or other. Whether they are fleeting observations, beliefs, or intense desires, they typically remain inaccessible to everyone but the individual with first-person access to them. Based on Nagel's observation, we can define *core private information* as an individual's unexpressed psychological properties to which only the individual has first-person access, and that are not knowable by anyone else, except by the individual's prior divulgence of them, or by an unreasonable inference based on other facts already known about the individual. *Core privacy* can be defined in relation to core private information as follows: A person, P, has core privacy in relation to a piece of core private information, I, and any other person, O, so long as I remains unexpressed by P, such that O can neither observe nor reasonably be expected to infer I.

The notion of reasonable inference is likely to raise objections, primarily because it may be considered too vague to underpin a theory of privacy. But in the case of data mining, what one can reasonably be expected to infer plays an important role in balancing claims to privacy. Generally speaking, any inference that an average unassisted person is capable of making given a set of data (about another individual) to which he has access via first-person observation, that is, a person who is not using a data mining algorithm, or working with a trained team of investigators, or referencing a database, etc., is a reasonable inference. This is intuitively compelling because by claiming that the gold standard for reasonable inference is an inference that could be made based on first-person observation, one maintains the social aspect of privacy as articulated by Rachels.[28] In addition,

[28] James Rachels, "Why Privacy is Important," (n. 11).

settling the question of reasonableness can be accomplished on a case-by-case basis—a requirement that mirrors the fundamental claims of the contextual approach.

Take our prediction about Jon's doughnut preference as a test case. Applying the synonymy test between the target property, namely Jon's desires regarding doughnuts, and our prediction, (i.e., an inference regarding Jon's doughnut preference) we would likely consider them synonymous. However, we would also conclude that it was reasonable to expect that any observer (one of his coworkers in this case) would have a good chance at inferring Jon's preference for chocolate doughnuts. So an application of the definition of core private information would lead us to conclude that Jon's doughnut preference was not, in fact, a private matter with respect to his colleagues. However, the inference might be unreasonable if it were made by an employee of Jon's credit card company, who would never be in a position to make the inference if not for her special access to vast quantities of data collected about Jon's and every other customer's purchases. So Jon's preference for chocolate doughnuts could be construed as core private information with respect to the employee, and Jon could have a reasonable claim to privacy with respect to that information. For cases in which the target predictions were less trivial, we would expect the determination of reasonableness to be less controversial.

Core privacy, as I have defined it, might strike some as too obvious to be worth discussing. *Of course* some of our unexpressed beliefs, intentions, and desires are private, maybe even "core" private; what of it? But if my description of emergent data is compelling, the current (and future) technological reality of data mining prompts the definition of core privacy; emergent data is nontrivial data obtained by O that has not been expressed by P, and that could not have been reasonably inferred by O. The corollary is that core privacy, once defined, can be used to assess the privacy implications of predictive data mining where a contextual analysis cannot. If the goals of predictive data mining are to produce emergent data that passes the synonymy test, then we need a methodology for assessing the privacy implications of those goals. Successful emergent data potentially contains core private information that *may not have been expressed by the individual in any context*. The upshot is that we have a way of describing how predictive data mining can violate core privacy.

Another way of interpreting the obviousness of core privacy is to say that it packs an intuitive punch, which goes a long way toward explaining peoples' intuitive objection to data mining. Explaining how predictive data mining can violate core privacy strikes a nerve in the same way that the fictional prospect of mind reading, as expressed by Nagel, would be intuitively problematic.[29] The use of far less credible technologies said to provide rudimentary mind-reading capabilities—fMRI scans touted as providing a visual means to predicting behavior or as

[29] Thomas Nagel, "Concealment and Exposure," (n. 27)

proving guilt in a crime, and polygraphs said to discern lies—has met with real criticism despite the limited success rates.[30] A concept like core privacy allows us to move beyond expressions of the intuitive unease we feel toward data mining practices, by providing a theoretical framework with which to identify and addresses some specific problems associated with the unique aspects of the technology.

VI. CONCLUSION

Predictive data mining differs from descriptive data mining in its use of KDD to discover nontrivial knowledge in datasets. KDD can produce emergent data when applied to datasets about individuals, the privacy implications of which contextual accounts of privacy fail to fully address. Assessing some of the unique privacy threats posed by predictive data mining can be helped by understanding the relationship between an individual's core private information and the predictions resulting from predictive data mining technology. If we are to better account for peoples' intuitive aversion to data mining, our focus should be on the emergent data—those properties that are meant to resemble an individual's beliefs, intentions, or desires. This shift in focus draws our attention to the unique potential that predictive data mining has for violating our core privacy. If "a *prima facie* case for caring about public surveillance is that it stirs popular indignation," then core privacy points to a *prima facie* argument that explains why that is the case.[31]

[30] J. Adler, "Mind Reading: The New Science of Decision Making. It's Not as Rational as You Think," *Newsweek*, August 9, 2004, http://www.newsweek.com/id/54762?tid=relatedcl.

[31] Nissenbaum, "Protecting Privacy in an Information Age: The Problem of Privacy in Public," 579 (n. 1).