Data Mining: An Annotated Bibliography

Solon Barocas[1]

Abstract

Data mining encompasses a diverse set of computational techniques that automate the process of discovering novel patterns and structure within large datasets. Such techniques are increasingly used to wrest actionable insights from the otherwise intractable datasets that often result from digitally mediated surveillance. For its relative obscurity, data mining has actually been the subject of significant inquiry across a diverse range of fields. This annotated bibliography surveys a selection of the canonical and emerging scholarship on the issues raised by the advent and various applications of data mining, reviewing the contributions of concerned technologists, legal scholars, philosophers, social theorists, and social scientists.

May 2011

[1] Doctoral Candidate, Department of Media, Culture, and Communication, New York University.

Introduction

For all the recent controversy, data mining is actually a well-established practice in marketing, advertising, and strategic communication, in consumer and corporate finance and banking, and perhaps most of all in insurance. It has also gained steady traction in policing and counterterrorism, in clinical and research medicine, and in all sorts of web-based services. And yet data mining remains shrouded in mystery.

The term itself—data mining—is the source of significant and chronic confusion. Because data mining often goes hand in hand with dateveillance, and because data mining nearly always lies at the foundation of automated systems, including those that engage in profiling, personalization, and targeting, much of the discourse tends to conflate these activities. 'Mining' suggests that datasets contain ready-formed nuggets awaiting discovery, but this is exactly wrong. Data mining doesn't discover already existing information; it generates new information about the existing data by revealing new and unexpected patterns and structure in the dataset.

Data mining helps to classify, cluster, or otherwise summarize the data; describe the dependencies, links, or trends within the data; and estimate or predict the value of future data (Frawley *et al.*; Fulda). These procedures differ from traditional data analysis in two important ways: (1) Unlike queries based on explicit criteria, data mining automatically generates the hypotheses it will then test. And (2) where information retrieval tools simply describe data that already exist in a database, data mining discovers patterns and structure that constitute new information about the data—information that can then serve as the basis for various forms of inference and automated decision-making. This can resemble a kind of 'mining' because it will often routinize the process of evaluating the significance of incoming data and thus help identify cases of the relevant activity in the subsequent flow of information.

These and other features of data mining have generated a surprisingly extensive and diverse literature outside its home fields. A fair amount of scholarship focuses on the nature of the analysis that data mining performs and the downstream impact of assumptions, bias, and errors on the populations subject to decisions informed by that analysis (Hilderbrandt and Gutwirth; Frawley *et al.*; Schauer; Fulda; Hilderbrandt; Vedder; Danna and Gandy; Custers; Poon; Canhoto). Other scholarship looks at the differential impacts of the kinds of activities and applications that data mining enables—that is, at the effects of putting data mining results to use in decision-making (Gandy; Lyon; Harcourt; Ford; Zarsky; Hilderbrandt; Danna and Gandy; Custers; Tverdek; Nielsen). Much of this work tends to focus on issues of privacy, justice, solidarity, autonomy, and individuality. Yet another cluster of research focuses on the difficulties that data mining poses to the fair information practice principles, including challenges to notice, consent, access, and transparency, but also other common mitigations, such as anonymization (Hilderbrandt and Gutwirth; Piatetsky-Shapiro *et al.*; Zarsky; Hilderbrandt; Schwartz; Tavani; Custers; Tverdek).

This annotated bibliography is divided into five sections. The first reviews some of the canonical critical literature; the second compliments this with a discussion of the seminal contributions and early concerns from the technical community. The next section considers two important interventions in the philosophy of law. This is followed by a more general review of the legal scholarship on commercial data mining. The document ten covers a few important publications in applied and practical ethics, before concluding with three recent research projects that indicate new directions for the study of data mining.

The Foundational Critical Literature

1. Gandy, Oscar H. *The Panoptic Sort: A Political Economy of Personal Information*. Boulder, CO: Westview, 1993.

Extending Foucault's work on panopticism to the information age, and drawing on social theorists and critics of bureaucracy, Gandy describes how "the collection, processing, and sharing of information about individuals and groups […] is used to coordinate and control their access to the goods and services." Gandy details a wide range of applications across business and government and describes profiling technologies as discriminatory by design—as tools that allow organizations to extend different options and opportunities to individuals and groups according to their estimated value and worth. Gandy stresses that such practices pose a threat to more than privacy; the discrimination they enable constrains life chances, exacerbates historical inequalities, and produces unequal access to information. Gandy goes on to show that the law offers little protection or recourse against these new dangers and, further, that such techniques may well undermine key aspects of democracy. The book's legacy has been to connect surveillance more directly to issue of fairness, equality, and distributive justice.

2. Lyon, David. *Surveillance As Social Sorting: Privacy, Risk, and Digital Discrimination*. New York, NY: Routledge, 2003.

Building on Gandy's 'panoptic sort' and observing that surveillance increasingly occurs to sort people into categories of worth and risk, Lyon and his colleagues, in this canonical edited volume, steer the discussion of surveillance away from its more common focus on privacy to issues of social justice. As in Gandy's work, discrimination and inequality are the clear focus: "The so-called digital divide is not merely a matter of access to information; information itself can be a means to creating divisions". Lyon is especially concerned with the criteria by which social sorting occurs: do the categories derive from prejudicial and stereotypical sources that simply reinforce or even exacerbate historical disparities? To get a better handle on the substance and significance of these criteria, Lyon advocates for a new research program that would look more carefully at the development of profiles, at the interaction between profiling technologies and their human operators, and at the experience of the populations subject to their evaluations and decisions. Many of the volume's contributors show that the categories and criteria of suspicion, eligibility, inclusion, and access can't be impartial because they emerge with specific (often commercial or state) purposes in mind.

3. Hildebrandt, Mireille, and Serge Gutwirth. *Profiling the European Citizen: Cross-Disciplinary Perspectives*. New York, NY: Springer, 2008.

The outcome of a highly interdisciplinary project on the Future of Identity in the Information Society, this edited volumes tackles three major questions: What is profiling is? Where is it applied? And what are its impacts on democracy and the rule of law?

Addressing the first question, Hilderbrandt advances a definition of profiling as pattern recognition. Canhoto and Backhouse describe the Knowledge Discovery in

Databases (KDD) process, focusing on the effects of technical, formal, and social norms on the development of profiles. Anrig, Browne, and Gasson describe the application of data mining algorithms to the datasets and the dynamics of the human-machine interaction. Yannopoulos, Andronikou, and Varvarigou explain behavioral biometric profiling, an approach to identification that relies, for example, on speech, facial features, and gait. Van der Hof and Prins looking into personalization.

The second section of the book reviews a number of real-world applications, including further cases of biometric profiling, location based services, the profiling of web users, attention support in work and school, and customer loyalty programs and credit scoring practices, among others.

To better address the third major question of the book, Schreurs, Hildebrandt, Kindt, and Vanfletern distinguish between the collection of data, the construction of group profiles, and the application of profiles to an individual person. They doubt that the Data Protection Directive, as it currently stands, can successfully protect against the 'unfair, illegitimate, or illegal' discrimination that might result from profiling; they also find the value of anti-discrimination law to be unclear. Gutwirth and De Hert note the difficulty of establishing transparency in the case of profiling because the profiles to which individuals are subject tend to come from an analysis of the data that belong to other people. They also question whether the trigger for data protection should remain personal information when many types of profiling are possible in the absence of such identifying information. Hilderbrandt considers the effects of profiling on personal identity and individual freedom. She shows how it affects both negative freedom by enhancing the capacity for institutions to interfere with individuals and positive freedom by limiting the choices available to individual.

The editors conclude that "a paradigm shift is needed from privacy and protection of personal data to discrimination and manipulation and transparency". In particular, they suggest that there is a need to better specify the legal status of profiles, to design systems that permit user experimentation and contestation, and to invent transparency-enhancing tools (TETs).

Seminal Contributions and Early Concerns from the Technical Community

4. Frawley, William J, Gregory Piatetsky-Shapiro, and Christopher J Matheus. "Knowledge Discovery in Databases: An Overview." *AI Magazine* 13, no. 3 (1992): 57-70.

In their attempt to consolidate a fledgling field, Frawley *et al.* proposed one of the canonical definitions of data mining: "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data". They coined the tern Knowledge Discovery in Databases (KDD) to refer to the overall process of transforming a question into a data mining problem, assembling and preparing the relevant data, subjecting the data to analysis (the data mining phase), and interpreting and implementing the results. The authors contrast KDD with traditional database retrieval and reporting techniques, with expert systems, with statistics, and with scientific discovery, finding in each case that KDD differs in often subtle but important ways. The article then considers a series of very practical concerns: how to deal with problems in the organization, completeness, quality, and certainty of the data; how to select and apply the right discovery algorithm; how to define and determine the certainty, accuracy, and interestingness of the discovered knowledge; how to integrate the user into the KDD process (by drawing on background and domain knowledge in setting out the problem and interpreting the result); and how to choose between the various forms that discovered knowledge can take. The paper offers a number of early success stories (from as early as the late seventies), noting other domains that are most likely to benefit from the application of KDD. The paper concludes with a generic framework that describes the full KDD process and the criteria that are likely to determine its success. Frawley et al. also note in passing that some discoveries might be illegal (in cases that involve protected classes) or unethical.

5. Piatetsky-Shapiro, Gregory. "Knowledge Discovery in Personal Data vs. Privacy: A Mini-Symposium." *IEEE Intelligent Systems* 10, no. 2 (1995): doi:10.1109/MIS.1995.10016.

   O'Leary, Daniel. "Some Privacy Issues in Knowledge Discovery: The OECD Personal Privacy Guidelines." *IEEE Intelligent Systems* 10, no. 2 (1995): doi:10.1109/64.395352.

   Bonorris, Steven. "Cautionary Notes for the Automated Processing of Data." *IEEE Intelligent Systems* 10, no. 2 (1995): doi:10.1109/MIS.1995.10015.

   Klösgen, Willi. "KDD: Public and Private Concerns." *IEEE Intelligent Systems* 10, no. 2 (1995): doi:10.1109/MIS.1995.10013.

   Khaw, Yew-Tuan, and Hing-Yan Lee. "Privacy & Knowledge Discovery: A Response to O'Leary." *IEEE Intelligent Systems* 10, no. 2 (1995): doi:10.1109/MIS.1995.10014.

   Ziarko, Wojciech. "Response to O'leary's Article." *IEEE Intelligent Systems* 10, no. 2 (1995): doi:10.1109/MIS.1995.10012.

In an early and overlooked exchange that predates the Data Protection Directive, a number of data mining scholars and practitioners identified the challenges that KDD posed to the principles advanced in the OECD Guidelines on the Protection of Privacy. The authors explain that, unlike traditional information retrieval, data mining actually returns *new* information (e.g., group profiles) that is not explicit in the original data, and that "such aggregate patterns are not covered by the restrictions on personal data" (Piatetsky-Shapiro). Klösgen argues that this is a worrisome gap in the law because analysts often apply the discovered group behavior to all of the members of the group. O'Leary therefore insists that "the KDD community must determine what kinds of data fall under the heading of personal". Piatetsky-Shapiro wonders whether the law could place limits on the kinds of patterns that data mining can discover without limiting the freedom of speech. O'Leary, in turn, asks whether data mining is even compatible with the principle of purpose specification: is it enough to explain that the data may be subject to mining or would the organization have to specify the patterns to be discovered (the value of which depends on their unexpectedness) and the purposes to which they may be put. He further questions the import of the openness principle, asking whether organizations would have to reveal the discovered knowledge and the process by which they arrived at that knowledge, even if it doesn't involve any personal information. The participation principle would also seem to suggest that data subjects should have the right to challenge the knowledge discoveries related or applied to them, but how would organizations implement such a systems? Piatetsky-Shapiro propose some possible remedies: anonymization, limited queries, aggregate statistics, and synthetic data.

Profiling and the Philosophy of Law

6. Schauer, Frederick F. *Profiles, Probabilities, and Stereotypes*. Cambridge, MA: Belknap Press of Harvard University, 2003.

Schauer sets out to challenge the "primacy of the particular"—the tendency to grant more legitimacy and moral weight to particularized judgments rather than judgments that rely on generalizations. He argues that generalizations are almost always unavoidable, that our cognition depends on extrapolating from like cases to a general rule, and that generalization "is one of the factors on which good judgment rests". The problem is how to generalize in a reasonable way.

Schauer distinguishes between "statistically sound generalizations"—generalizations for which there is good statistical evidence—and spurious generalizations. Schauer explains that most cases fall somewhere between these two extremes. He describes these as "statistically sound, but nonuniversal generalizations"—generalizations that apply to some, but not all individuals who conform to the profile that is meant to predict a certain quality or propensity. Schauer's book is essentially an extended discussion of when the use of statistically sound, but nonuniversal generalizations is justified.

Schauer argues that where the costs or barriers to direct observation of the behavior of interest are sufficiently high, profiling may be reasonable, even when such generalizing may involve occasional errors. He argues that the attempt to avoid profiling might actually affect or inhibit decision-making in ways that introduce their own costs. That said, Schauer notably opposes racial profiling at the airport because he believes the salience of race would lead to its overuse in rendering decisions. He also points out that the costs of the policy would fall disproportionately on the targeted population, which would put a strain on race relations and contribute to stigmatization. He concludes that it would be preferable to improve security by subjecting all passengers to increased scrutiny—adding time to everyone's journey—rather than singling out specific groups.

7. Harcourt, Bernard E. *Against Prediction: Profiling, Policing, and Punishing in An Actuarial Age*. Chicago, IL: University of Chicago Press, 2007.

Harcourt advances a three-pronged critique of prediction and profiling in the service of policing, sentencing, and parole decisions. He first shows how the advent of actuarial techniques in policing tends to substitute a concern with crime reduction (efficacy) with a narrower concern for whether each police action results in crime detection (efficiency). Harcourt argues that such an approach to policing may actually result in *more* crime:

while subjecting a subpopulation that has demonstrated higher offense rates to increased scrutiny may well result in an increase in detected crime (i.e., each individual search is more likely to result in an arrest), it may simultaneously increase the overall amount of crime because the reallocation of resources means that a larger relative proportion of the offenders from the rest of populations go unnoticed. Harcourt further points out that because different populations are likely to differ in their responsiveness to policing, it's quite possible that the more successful targeting of policing may simply increase the discovery of crime while doing little to actually reduce it. This bleeds into Harcourt's second point: the targeted subpopulations will begin to account for a larger relative proportion of the imprisoned population—larger than their representation among actual offenders—and this will place increased stress on the communities to which these offenders belong. Harcourt therefore concludes that profiling contributes to a ratchet effect that only reinforces many of the social costs associated with crime. And this, in turn, further fuels stratification and stigma. Finally, Harcourt argues that profiling violates the basic idea that "anyone who is committing the same crime should face the same likelihood of being punished," and that randomization is the only way to hold true to this principle.

The Legal Scholarship on Commercial Data Mining

8. Fulda, Joseph S. "Data Mining and Privacy." *Albany Law Journal of Science and Technology* 11 (2000): 105-113.

Fulda sets out to determine whether data mining constitutes a violation of privacy, and whether tort law and the Fourth Amendment offer sufficient protection against possible violations. Drawing on the technical literature, Fulda first distinguishes data mining from informational retrieval: unlike traditional database queries that simply return the records that match the query criteria, data mining instead reveals latent patterns in the data and thus results in a kind of derived or inferred knowledge. He explains that data mining can classify, cluster, or summarize the data; describe the dependencies, links, or trends in the data; and predict the value of future data. He therefore asks, "Is it possible for data that does not in itself deserve legal protection to contain implicit knowledge that does deserve legal protection"? Fulda works through a series of hypothetical situations that involve similar modes of inference as data mining to see whether such protections would be warranted. In each case, an observer takes two pieces of otherwise public information to infer and make explicit a third piece of previously private information. If such cases constitute a violation, as Fulda argues, then so, too, should data mining: "if data about an individual is mined and implicit knowledge about him is discovered, an appropriation has occurred, and further disclosure should not be permitted". Fulda notes, however, that case law does not support this position and therefore fails to offer sufficient protection. Fulda concludes by suggesting cryptographic solutions as an alternative defense.

9. Ford, Richard T. "Save the Robots: Cyber Profiling and Your So-Called Life." *Stanford Law Review* 52, no. 5 (2000): 1573-1584.

In a prescient thought experiment, Ford wonders what would happen if we increasingly come to rely on the suggestions of recommender systems based on user profiling and data mining (e.g., collaborative filtering). He describes a 'cyber doppelgänger': the profile that recommender systems use to predict our next move and future interests. If the resulting recommendations are good enough, we are likely to rely more and more heavily of the selections of the cyber doppelgänger and less on our own judgement and the limited information we can gather. Moreover, our cyber doppelgänger will offer more considered choices, helping to counteract our tendency to make poor, irrational, or impulsive choices. It will know us better than we know ourselves and it will allow us to be more true to ourselves: "If we have given up autonomy, it was only the autonomy to make poor choices, go to bad restaurants with people we turn out not to like much, buy boring novels, listen to ear splitting music, engage in activities where costs outweigh benefits. I am actually more free now than ever before because my true self—the self that labored under misconceptions, limited information and emotional noise—is now facilitated by powerful and benevolent technology".

Ford disputes some common concerns: companies are unlikely to exploit our dependency on recommender systems by charging us for the privilege of their advice; in acting on those suggestions, we serve businesses' interests. Recommender systems are unlikely to fall pray to the digital divide; although they will recommend different things to differently situated people, recommender system will have as much advice to offer the poor as the rich. Finally, concerns around autonomy seem misplaced because the recommendations are merely suggestions; they don't constrain choice or information-seeking. He acknowledges that there might be financial incentives and social pressures to accept the recommendations, like discounts on insurance, for example. But Ford's main concern is more profound: as we happily and increasingly turn to recommender systems for advice, those systems begin to shape our own thoughts and desires. In predicting more and more of our interests, the recommender systems will start to change them. Paradoxically, then, we lose control of our own desires by relying on a system that attempts to cater to them. Ford suggests that we raise awareness about such possible futures so we can act before they arrive.

10. Zarsky, Tal Z. "Desperately Seeking Solutions: Using Implementation-Based Solutions for the Troubles of Information Privacy in the Age of Data Mining and the Internet Society." *Maine Law Review* 56, no. 1 (2004): 13-59.

To better understand the problems that result from the advent of commercial data mining, and to better locate the stages at which these problems arise, Zarksy breaks the data mining process into three separate parts. The fist part is the collection phase, where information is initially obtained; the second part is the analysis phase, where the actual data mining take places; the third and final part is the implementation phase, where the results of data mining are put to use. Zarsky argues that most solutions to the problems introduced by data mining tend to focus on the collection stage, even though such solution are less effective and possibly less fair than alternatives that focus on the implementation phase.

Solution that limit collection tend to run up against the freedom of speech. Negotiated contracts, based on property rights, would involve enormous transaction costs that severely disadvantage the consumer. In general, it is exceedingly difficult to provide upfront and complete notice in the case of data mining because the value and potential uses of any piece of data are often impossible to predict. Blanket prohibitions on secondary use, at the analysis stage, would severely limit many positive uses of the data, including improved resource allocation (especially in marketing)—improvements that tend to benefit everyone because they reduces waste and thus reduces prices. Protected classes, prohibiting the use of certain variables, might be a more effective solution.

Zarsky then shows how the problems that may arise with data mining—manipulation and threats to autonomy; abuse and misuse; intrusion on seclusion; adverse effects of errors; discrimination—are best addressed by focusing on the implementation phase. For example, Zarsky supports a right of access to deal with the harms that result from inaccurate data or inferences, though he suggests that this right be limited to the personal information pertaining to the individual and not include access to the patterns and categorization that result from data mining. To limit the possible manipulation of choices through the use of personally targeted content, Zarsky recommends a notice and diversity requirement. And to address price discrimination, Zarsky suggests the use of tools to facilitate communication between consumers, thus ensuring that possible secondary markets discourage vendors from attempting to carry out such schemes. He also discusses medical and life insurance as a unique case of price discrimination, noting the problems with adverse selection. Zarksy argues that blocking the collection of information to ensure the socialization of risk will lead to arbitrary subsidies for certain policyholders who happen to benefit from the current flow of information that nonetheless continues to allow companies to differentiate between customers. He argues that solutions should confront the matter directly at the implementation stage by having the government set maximum insurance rates or by prohibiting the consideration of certain factors (e.g., gender or disability). Zarsky concludes that better solutions are those with less distance between the point of regulation and the actual use of information.

11. Hildebrandt, Mireille. “Who Is Profiling Who? Invisible Visibility.” In *Reinventing Data Protection?* Edited by Serge Gutwirth, Yves Poullet, Paul de Hert, Cécile de Terwangne and Sjaak Nouwt. Berlin, Germany: Springer, 2009, 239-252

Hildebrandt argues that there are two main hazards that stem from profiling. The first concerns the increased capacity of institutions that make use of profiling to engage in discrimination, particularly differential treatment that affects individuals’ autonomy. The second concerns the increased information asymmetry between the profiler and profiled, such that the profiled may know less about themselves than the profilers.

Because profiling allows organization to infer potentially sensitive facts form trivial and anonymous data, and because the profiles that are applied to an individual are derived from *other* people’s data, consumers are generally unable to determine the full consequences of their decision to disclose certain information. Hildebrandt concludes that informed consent is “wholly inadequate” in the case of profiling because the full implications of disclosure are only visible to those who engage in profiling, not those who are profiled. This lack of knowledge about the profiles opens people up to manipulation, but it also facilitates price discrimination, actuarial justice, and social sorting in ways that run counter to the principles of non-discrimination and due process.

Hildebrandt argues that the legal status of profiles, according to the Data Protection Directive, is quite ambiguous. First, the decisions taken on the basis of profiling are rarely fully automated, and thus not covered by Article 15. And even in the event that they are fully automated, these decisions may be covered by the caveat in Article 15 that allows for certain lawful applications. Hildebrandt points out that even if such protections applied, individuals would have to know that they are subject to profiling before they could resist the automated decisions taken on that basis.

Hildebrandt concludes by proposes an effective right of access to profiles, including access to information that explains the consequences of their application, as well as a right to contest the accuracy and applicability of a profile. She also proposes a correlate obligation on the organization that engage in profiling to communicate the profiles that match a person's data—a feature that should be built into the technological infrastructure of the profiling system itself. Hildebrandt ultimately advocates for a principle of 'minimum knowledge asymmetry.'

12. Schwartz, Paul M. "Data Protection Law and the Ethical Use of Analytics." Richmond, VA: The Centre for Information Policy Leadership, 2010.

Drawing upon interviews with practitioners from some of the companies most well known for their use of analytics, Schwartz sets out to establish a set of legal and ethical guidelines that other organizations might follow when they employ analytics. Schwartz adopts a broad definition of analytics: tools that "[h]arness statistics, algorithms, and other tools of mathematics to improve decision-making". He argues that different applications of such tools are likely to present very different regulatory and ethical issues. Working through a series of real-world examples, he shows how marketing (e.g., online advertising), fraud prevention, healthcare research, and products for direct use by individuals (e.g., recommender systems) don't present the same set of issues. The issues tend to depend on the categories of business, which leads Schwartz to suggest the need for a 'contextual analysis' of analytics that would evaluate the different kinds of risks that arise in different applications.

Schwartz nonetheless offers a set of general prescriptions: (1) comply with the law; (2) assess whether the use of analytics is in keeping with cultural and social norms; (3) determine stakeholder impact; (4) develop accountability procedures with clear assignments of responsibility; (5) ensure the security of data, proportionate to the sensitivity of the information; (6) determine whether the use of analytics involves sensitive areas and put in place reasonable safeguards proportionate to the risk; and (7) consider the special vulnerability of children. He also offers a set of recommendations for each stage in the analytics life cycle—a cycle that he breaks apart into (1) collection, (2) integration and analysis, (3) decision-making, and (4) review. He argues that even when analytics thrive on access to lots of data, companies should refrain from collecting certain information. He also recommends that companies exclude certain data from the analysis, especially when that data adds little to overall efficacy of the effort, and that they anonymize personal information when appropriate. He then goes on to suggest that the results of analytics should be reasonably accurate in proportion to the nature and significance of the decisions they will inform. Aside from providing notice, access, and other remedies, including controls for users to express preferences, companies should assess whether the decisions they take on the basis of analytics violate prevailing legal, cultural, and social norms. Finally, companies should constantly review whether the information that figures in the analytic process continues to retain its predictive value and whether the results of analytics actually improve decisions.

Applied and Practical Ethics

13. Tavani, Herman T. "KDD, Data Mining, and the Challenge for Normative Privacy." *Ethics and Information Technology* 1, no. 4 (1999): doi:10.1023/A:1010051717305.

Tavani sets out to show how data mining differs from traditional information retrieval techniques and why these differences present serious challenges to normative theories of privacy, privacy law, and the fair information practices. Tavani explains that data mining unearths patterns and relationships that can then be used to reveal information about individuals that is not explicit in the original data. Tavani further explains that data mining involves open-ended queries in which "the data is 'sifted' in search of frequently occurring patterns, trends, and generalizations about the data without intervention or guidance from the user". He contrasts this 'discovery' approach to more 'traditional' information retrieval techniques in which only those records that match the user-specified search criteria are returned. The open-endedness of data mining means that the value and potential uses of the discovered information are very hard to predict—an aspect of data mining that runs up against the principles of purpose specification and use limitation. Tavani concludes that "individuals cannot possibly be told in advance what kind of information data mining algorithms will yield about them and how that information will be used". Tavani also points out the new categories of persons that result from data mining implicate everyone's data because they derive from the data themselves, not just the data of the individual to whom these categories are applied. Current laws and theories of privacy fail to account for such situations. Tavani rightly anticipated that many of these problems would intensify as the Internet became a more common source of data, including biographical information made available by users themselves (e.g., on social networking sites).

14. Vedder, Anton. "KDD: The Challenge to Individualism." *Ethics and Information Technology* 1, no. 4 (1999): doi:10.1023/A:1010016102284.

Vedder is principally concerned with the 'deivididualization of the person': the tendency to judge and treat people on the basis of the groups profiles that result from data mining rather than people's individual qualities. Vedder worries especially about what he calls 'non-distributive group profiles': cases where an individual fulfills all the criteria for inclusion in a particular group, but fails to poses the quality that these criteria are expected to predict. Because data mining is often used to predict group membership, individuals may find that they have been placed into groups that do not accurately reflect their actual status.

Vedder explains that these group profiles do not constitute personal data and are therefore not subject to data protection, even though they are applied to individuals *as if* they were personal information. Vedder says that such applications should raise concerns not only because they may rely on potentially erroneous categorizations, but also because even accurate categorization can influence allocation decisions that might well give rise to discrimination, social stratification, and stigmatization. Vedder argues that existing privacy law and ethical theories concerning privacy fail to adequately capture these problems because they rely on too narrow an understanding of personal data.

After demonstrating the futility of responses that depend on a right of collective privacy (there are no ties of loyalty among members of the group that would provide a basis for collective action) and non-participation (individuals may refuse to contribute information to the profile-building exercise, but nonetheless remain subject to it), Vedder advances a notion of categorical privacy that calls for limits on the application of group profiles to individuals. He argues that a case-by-case assessment may be needed to determine when to restrict the use of profiling for certain purposes, and that such assessments should draw on additional normative principles other than privacy, including social justice, equality, and fairness.

15. Danna, Anthony, and Oscar H Gandy. "All That Glitters Is Not Gold: Digging Beneath the Surface of Data Mining." *Journal of Business Ethics* 40, no. 4 (2002): doi:10.1023/A:1020845814009.

Danna and Gandy show how the "objectively rational business decisions" that result from data mining can nonetheless contribute to a number of social costs. The authors identify three different types of costs: the first concerns the use of data mining to exclude classes of consumers from full participation in the market place; the second stems from the use of data mining to limit customers' or citizens' access to information essential to their full participation in the public sphere; and the third springs from the different kinds of errors that invariably enter into or result from the data mining process.

As Danna and Gandy show, profiling and personalization can make possible various forms of price discrimination, ranging from selective discounts and inducements for more profitable customers to differential pricing systems that attempt to predict the maximum price that a particular customer would be willing to pay for a product or service. Redlining is also possible: firms may decline to further service customers who have proven less profitable or they may simply refuse to service specific classes of customers outright. Profiling and personalization may also exacerbate inequalities in exposure and access to information. By limiting solicitations to more highly prized classes of consumers, firms may undermine other consumers' ability to make fully informed choices in the marketplace or to participate effectively within the public sphere. This is especially worrisome in the cases of policy-related information, where changes in the flow of information may distort public discussion and debate.

Dana and Gandy also stress the "consequences that flow from the accumulation of […] errors" in data mining process. These begin with the data themselves, which are vulnerable to a host of collection, collation, and coding errors. Profiling will invariably result in some inaccurate predictions, such that individuals may be assigned to groups to which they don't actually belong. Commercial firms are generally willing to accept a certain error rate because the benefits that accrue from accurate profiling far outweigh what they view as the costs. Even when profiling is accurate, it doesn't capture the full complexity of the profiled individual; people are treated as members of groups instead of distinct individual. Finally, because profiling draws from the past to predict the future, it necessarily excludes "human serendipity" from the model and fails to account for future circumstances that might change an individual's behavior.

16. Custers, Bart. *The Power of Knowledge: Ethical, Legal and Technological Aspects of Data Mining and Group Profiling in Epidemiology*. Nijmegen, Netherlands: Wolf Legal Publishers, 2004.

Focused on profiling in the context of epidemiology, Custer's *Power of Knowledge* examines the technical aspects of the data mining process to better understand its implications for individual patients, clinicians, insurance companies, hospitals, the government, and society, generally. Custer shows how data mining differs from traditional data analysis in that data mining automatically generates the hypotheses it will then test. Further, Custer explains that the discovered correlations are not necessarily casual, but that there are many cases, including epidemiology, where underlying causes do not have to be known in order to act upon the identified statistical relationships. Custer highlights the risks and effects of different inferential errors. Like Schauer and Vedder, Custers distinguishes between distributive and non-distributive profiles. He then goes on to show how the reliability of group profiles is an important factor in determining the positive and negative effects that the group profile may have.

Even where data mining and profiling offer clear benefits by allocating (often scarce) resources more efficiently through targeting and customization, they also pose certain dangers. In particular, he shows how profiling implicate: justice (by specifying the criteria for equal or unequal treatment); solidarity (by individualizing risks and effecting the distribution of opportunities through adverse selection and cream skimming); autonomy (by selectively influencing or limiting individuals' choices of decision-making); individuality (by relying on a limited number of predictive factors to describe and categorize unique individuals); honesty (through misleading consent agreements that cannot anticipate the unpredictable results of data mining and their many possible uses, but also through masking (i.e., the substitution of seemingly trivial factors for sensitive factors, where those trivial factors directly correlate with—and thus act as a proxy for—the sensitive factors)); and trust (by jeopardizing the confidence that patients place in their doctors).

Custers attempts to find concrete solutions to these problems by looking to the law, to technology, and to norms. He considers Dutch data protection law, public health law, and anti-discrimination law, finding that each have their limitations: data protection only covers cases that involve personal information (data mining can work even if the specific person to whom the data refers is not identifiable by name); the Dutch 'right not to know' (about one's possible medical condition) only regulates the release and collection of data, not the discoveries that come from analysis; and anti-discrimination statutes only cover cases where an actual selection took place, not the analysis that informed the selection decision. A general lack of transparency and the possibility of masking also mean that it would be exceedingly difficult to prove discriminatory selection. Common technical solutions also offer little help. Anonymity doesn't prevent data mining or group profiling; algorithms and targeting schemes work equally well with nameless database entires. The only way to ensure that data subjects do not open themselves to the application of the resulting profiles would be to make a record unlikable to any *particular* person. Of course, this would also make many other (and less worrisome) analyses impossible. That said, access and inference controls could limit unauthorized or unnecessary disclosure.

Custers ends with a number of recommendations. He suggests that data protection expand to cover not just personal data but also group profiles. Data protection law should also be more strictly enforced, forcing organizations to be more transparent about—and thus accountable for—their profiling practices. Anti-discrimination law should also reverse the burden of proof in cases where unequal treatment is suspected; the organization that rendered the decision in dispute should have to prove their innocence. He further explains that medical institutions, and insurance companies in particular, should separate the tasks of planning from insuring. Finally, because many of these recommendation rest of the willingness of certain stakeholders to consider the vulnerabilities of other stakeholders, Custers advocates a concerted attempt to improve awareness of the issues around data mining and profiling through education. A mix of legal, technological, and bureaucratic solutions appropriate to the context of use could then significantly reduce the dangers associated with data mining.

17. Tverdek, Edward. "Data Mining and the Privatization of Accountability." *Public Affairs Quarterly* 20, no. 1 (2006): 67-94.

Tverdek begins by arguing that the concern with government data mining "miscasts the privacy risks we actually face with the widespread use of data mining" because it occludes the far more numerous and consequential applications in the commercial sector. Tverdek also takes issues with the notion that commercial data mining is more legitimate because it relies on data volunteered by individuals, unlike the data that the state might compel its citizens to disclose. He points out that many of the choices that individuals face in their interaction with commercial enterprise are not really choices at all because the costs of non-participation are often profound. Even more troubling, however, is the fact that individuals are increasingly held accountable for their actions with information that they surrendered for other, often benign or beneficial use. Tverdek's main concern with data mining lies in private firms' capacity to draw inferences that better equip them to hold individuals accountable for their actions, even though these very same organization are not subject to democratic controls.

Tverdek further points out that the capacity of private firms to hold any particular individual accountable depends on the willingness of *other* individuals to reveal data about themselves. Taking the example of car insurance, Tverdek explains that what constitutes a 'good' driver is a statistical artifact that can only be made in contrast to a statistically 'reckless' driver; 'good' drivers will have incentives to volunteer information that will lower their rates, but they will also have incentives to make the sharing of this data mandatory for even those who would not wish to volunteer it. Ultimately, Tverdek concludes, "as long as some significant number of people opt in to such a regime, non-participants and the statistically risky are held accountable in the same way".

Tverdek argues that such conditions call for democratic institutions with protocols for collective decision-making. In the absence of such protocols, the information revealed by other people will continue to allow private companies to determine whether individuals are “indolent, accident-prone, or squanderous,” even though those who might object to such determinations have few, if any, means to exclude themselves from—or directly challenge the judgment of—the system. Tverdek points out that the democratic processes has resolved such tensions in the past; in the case of the Equal Credit Opportunity Act, Congress forbid the use of certain data to draw inferences about an individual’s character and prospects.

Tverdek concludes with a discussion of ascribed and achieve characteristics and the socialization of risk, challenging the notion that just because individuals engage in activities that they could forgo if they chose, that they somehow surrender a part of their autonomy, granting consent to the use of the fact of that activity or those characteristics. Tverdek further argues that the voluntary and involuntary nature of the activity seems to be entirely independent of the moral question of how much control individuals might justifiably retain over who knows they do it and how private companies may use this information.

New Directions

18. Poon, Martha. “What Lenders See: A Pre-History of the FICO® Credit Scoring System." Unpublished Dissertation. San Diego, CA: University of California, San Diego, 2010.

Drawing upon insights from the social studies of science, technology, and finance, Poon develops a history of credit scoring that traces the early origins of data mining, through operations research, in consumer finance. Her work demonstrates the enormous difficulty Fair Isaac faced in trying to convince the industry of the utility of its scoring technology. Drawing from trade literature and government records as well as interviews with Fair Isaac employees, she details the various steps that Fair Isaac took in the post-war period to make its scores meaningful and useful in terms there were already available and relevant to the industry and government. Poon reveals that the company’s algorithms were designed to overcome a subtle but important tension between an earlier propensity for fixed classification schemes and post-war approaches to managing business operations through probabilistic mechanisms. She concludes that credit scoring succeeded because of its repeated technical adaptation rather than its ability to assess risk in formal scientific terms, and that the course of these adaptations opens the U.S. credit scoring system to further political intervention.

19. Canhoto, Ana Isabel. “Profiling Behaviour: The Social Construction of Categories in the Detection of Financial Crime.” Unpublished Dissertation. London, UK: London School of Economics and Political Science, 2007.

Taking the use of data mining in an anti-money laundering (AML) program in the UK as her case study, Canhoto examines the mutual influence that profiling technology and their users have on each other and the resulting behavioral profiles. Anti-money laundering presents a unique case because, unlike most other instances of data mining, there are only a limited number of documented cases of financial crime from which to infer a profile. Profiling in the case of AML therefore involves a complex mix of automation and ongoing interpretation and judgment. She draws on organizational semiotics and classification theory (form cognitive science) to build a model of the technical, formal, and informal factors that influence the ultimate shape of the profiles. While her work stresses the ability and need for human agents to exercise meaningful influence on the development process, the results of her fieldwork also show that profiles are really the result of a dialectical process: the data and analytic techniques structure the analysts' interpretations at the same time that the analysts tend to search the data for only those things which the analysts already believe to be reasonable and conceivable. Canhoto describes this as a 'sequential meaning-making process', and she highlight the different cues that direct its movement. Understanding this process allows Cahoto to say when and where bias and undesirable influences can enter the process and how to better contain them.

20. Nielsen, Rasmus Kleis. *Ground Wars: Personalized Communication in Political Campaigns*. Princeton, NJ: Princeton University Press, 2012.

Rasmus Kleis Nielsen has looked at a very different part of the data mining process: executing a communication strategy based on data mining results. Nielsen offers an ethnographic account of two congressional campaigns in 2008 that reveals how new information technologies have enhanced—rather than replaced—old forms of political communication. Responding to oversaturation and audience fragmentation in mass media, campaigns have renewed their interest and investment in 'ground war' practices like canvassing and phone banking as a way to better communicate with voters. Nielsen specifically focuses on the increased precision with which ground campaigns are now conducted because of the advent of data mining—helping to better locate receptive voters and better predict the issues they care about—but he also details the many challenges that campaign managers face in trying to get volunteers and canvassers to stick to the targeting strategy that data mining suggests. There is a cruel irony to Nielsen's story: the individual volunteers who actively participate in the political process are precisely those who are stripped of their ability to exercise independent political opinions or to engage in spontaneous political debate. They now instead act as the vehicles for precise, pre-scripted messages.