



# The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making

Science, Technology, & Human Values

2016, Vol. 41(1) 118-132

© The Author(s) 2015

Reprints and permission:

[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)

DOI: 10.1177/0162243915605575

[sthv.sagepub.com](http://sthv.sagepub.com)



Tal Zarsky<sup>1</sup>

## Abstract

We are currently witnessing a sharp rise in the use of algorithmic decision-making tools. In these instances, a new wave of policy concerns is set forth. This article strives to map out these issues, separating the wheat from the chaff. It aims to provide policy makers and scholars with a comprehensive framework for approaching these thorny issues in their various capacities. To achieve this objective, this article focuses its attention on a general analytical framework, which will be applied to a specific subset of the overall discussion. The analytical framework will reduce the discussion to two dimensions, every one of which addressing two central elements. These four factors call for a distinct discussion, which is at times absent in the existing literature. The two dimensions are (1) the specific and novel problems the process assumedly generates and (2) the specific attributes which exacerbate them. While the problems are articulated in a variety of

---

<sup>1</sup>University of Haifa, Haifa, Israel

## Corresponding Author:

Tal Zarsky, University of Haifa, Mount Carmel, Haifa 31905, Israel.

Email: [tzarsky@law.haifa.ac.il](mailto:tzarsky@law.haifa.ac.il)

ways, they most likely could be reduced to two broad categories: efficiency and fairness-based concerns. In the context of this discussion, such problems are usually linked to two salient attributes the algorithmic processes feature—its opaque and automated nature.

### Keywords

privacy, big data, automatic decisions, discrimination, data protection, credit scoring

## Strategy and Motivation

As recent literature and popular news reports demonstrate, the use of algorithmic decision-making tools by governments and private entities has grown progressively. Algorithms are applied to various forms of data and in numerous settings, often relying on the analysis of personal information. In light of this change, a new wave of policy concerns has emerged (Pasquale 2015; Citron and Pasquale 2014; Zarsky 2013a; Crawford and Schultz 2014). This article broadly maps these concerns, aiming to “separate the wheat from the chaff” and to provide policy makers and scholars with a comprehensive framework and taxonomy to approach these thorny issues. In the interest of simplicity, this article focuses on a general analytical framework to explain the troubles with algorithmic decision making, yet at the same time applies it to a specific subset and test case.

Providing an overall analytical framework for the vast topic of algorithmic decision making is an overambitious task. Therefore, this article reduces the discussion to two dimensions: (1) the specific and novel *problems* that algorithmic decision-making processes generate and (2) the *attributes* that exacerbate these problems.

Furthermore, both dimensions must be somewhat narrowed down. Therefore, while the *problems* arising during algorithmic analysis of personal data are varied, two broad categories can be carved out: *efficiency*- and *fairness*-based concerns (see Perry and Zarsky 2014 for a similar methodology in other contexts). Similarly, the article focuses on two salient *attributes* of the algorithmic decision-making process, its *opacity* and its *automated* nature: algorithmic decision making often uses nontransparent measures and analyzes data automatically. To summarize, Table 1 displays the analytical dimensions this article will explore. Actions taken to resolve one concern may not solve the other and may even exacerbate it.

**Table 1.** Governing Algorithms: An Analytical Framework.

<b>Problems</b>	Efficiency	Fairness
<b>Attributes</b>	Automation	Opacity

This article focuses on credit ratings (or “scores”) used by financial institutions to determine whether to offer or deny credit to individuals (for a similar analytical choice, see Citron and Pasquale 2014; Zarsky 2014). Credit assessment is to a large extent carried out by an automated and opaque process. This process, relatively unique to the United States, has replaced a subjective analysis that used to be carried out by local credit officers. Many would argue that adopting this scoring process has led to vast benefits, such as lowering the cost of credit. Yet, others claim that these benefits have high costs and cause a variety of problems and concerns (Citron and Pasquale 2014). Therefore, a focus on credit assessment and scoring allows us to tap into relevant discussions about algorithmic governance in academia and the public policy sphere, though here I do not offer concrete policy recommendations regarding this matter.

Before proceeding, I want to note three potential objections to the proposed methodology: *First*, the discussion addresses conceivable complications resulting from automated and opaque processes. Yet, it does so while merely acknowledging (rather than fully discussing) two central and intuitive arguments—that (a) automation is categorically problematic and (b) transparency is categorically important. Both arguments have been made in the literature. Nonetheless, this article strives to reach beyond these supported claims by pointing to second-order considerations that allow us to judge the benefits and detriments of both automation and transparency. In addition, reaching beyond the categorical arguments noted above permits calibrating the role that automation and transparency should have in specific social contexts.<sup>1</sup>

*Second*, any institutional decision that applies or allows algorithms to automatically sort, govern, and decide issues related to human actions makes two crucial assumptions: that human conduct is consistent and that with sufficient data human behavior becomes predictable. Suffice it to say, the notion that human behavior is somewhat predicable may be a difficult concept to accept. This assumption engages an almost infinite literature in the fields of philosophy, biology, psychology, theology, and others—clearly a task beyond the confines of this article. Thus, this article sidesteps a discussion on the predictability of human nature and accepts it, while

recognizing the limited errors it might entail, on the aggregated level (when predicting aggregated, as opposed to specific, human behavior).

*Third*, the following analysis initially addresses automation and transparency as separate elements. However, the elements impact each other and there are interesting interactions between the two.<sup>2</sup> For instance, a high level of automation in algorithmic processes could inherently increase opacity. Analysis based upon mined data, premised on thousands of parameters, may be difficult to explain to humans. Therefore, achieving transparency in such cases presents substantial challenges. Equally, the firm governing through such data analysis would find it difficult to adequately explain the “real reason” for its automated response—even after making a good faith effort to do so. Similarly, meeting transparency requirements might require limiting automation.<sup>3</sup>

Another point connecting automation and transparency concerns human deference to machine-driven decisions. People tend to accept that automated processes are true and precise (Citron 2007). Humans, some argue, will happily defer to the machine. Yet such blind deference is ill founded. For this reason, greater automation calls for greater transparency—albeit, specific forms of transparency to assure that the public understands that automated processes are also fallible. Clearly, transparency will only be one part of an overall effort in public education on the true nature and possible shortcomings of algorithmic governance.

## **Governing Algorithms: The (In)Efficiency Argument**

A reasonable and popular critique of the algorithmic process is that it is inefficient, decreases welfare, and, perhaps, even destroys value. Key elements of the efficiency-based argument note several distinct sets of flaws. First, inaccuracies in the underlying data sets could hamper the process. Second, the analysis itself could be mistaken due to errors in predicting individual behavior, since, in specific settings (and counter to the overall assumption noted in the introduction), human conduct is unpredictable. These errors undermine the supposed benefits of the algorithmic process. In the context of credit scores, for example, the “inefficiency” argument assumes that this error-prone process will erroneously calculate individual risks to default on loans. Therefore, the assumed benefits of lowering credit rates for the entire public, such as overall improved access to credit (given the lender’s lower costs), may not materialize. Yet, if even a problematic algorithmic process provides overall benefits that render it better than

alternatives (and with similar costs or externalities<sup>4</sup>), this efficiency-based critique of governing via algorithm fails.

While errors in both the data and its processing exist, the scoring process can still yield an accurate credit level to applicants on an aggregated level. Various errors might offset one another, are corrected over time, or have only a limited effect on the final outcome. In other words, merely pointing to anecdotal errors is insufficient, and a more extensive analysis of relevant markets is necessary. Yet those voicing the critique often merely point to the specific errors revealed, rather than to other overall factors of higher interest rates or credit costs caused by the system's inefficiency.

And what are the alternatives?<sup>5</sup> Automated algorithmic processes that assess creditworthiness err at times, but manual credit rating and sorting can err too. Nonautomated processes spawn a separate set of problems related to human judgment. Humans tend to make systematic and predictable mistakes, and our decisions are subject to bias (Jolls, Sunstein, and Thaler 1998). In this sense, underlying data and processing errors that characterize algorithmic credit scoring processes do not in themselves sufficiently demonstrate the relative benefits of manual credit scoring techniques. Quite to the contrary, these errors of human judgment and bias might be mitigated in the automated environment (Meadow and Sunstein 2001).

Transparency arguably may correct errors in any algorithmic process, thus promoting efficiency. It allows individuals to correct inaccurate *data* that have been collected about them. In this way, transparency also brings the scrutiny that will pressure agencies to improve their practices. This leads to the conclusion that transparent processes will prove more accurate and, thus, efficient.

Yet, reality calls into question the adequacy of this pro-transparency argument. For example, individuals have a right to review and correct credit records (as well as many other personal data sets, such as health records), yet very few do so (Hunt 2005). The public (and the media) usually shies away from the close analysis of the technical mechanisms of algorithmic analysis that such disclosures might require (Lenard and Rubin 2013). Even if transparency somewhat improved the accuracy of algorithmic processes, the aggregated costs of facilitating disclosure (and the losses that mount as a result of public scrutiny) render it costly. Once we acknowledge such factors, transparency does not appear to substantially enhance social welfare.

Indeed, the algorithmic credit scoring system strives to predict future impermissible behavior (such as defaulting on loans) while relying upon a set of behavioral proxies. If transparency allows identification of behavioral indicators of credit risk, individuals will try to avoid being linked

to these behaviors and indicators. Yet, the overall negative outcome of individual behavior need not change. In other words, with full transparency, monitored individuals will sidestep proxies even while still engaging in risk-generating behavior. For instance, they might refrain from using their credit cards at discount stores (a possible negative proxy) but continue to spend in general. Additional study must follow to establish whether this problematic outcome is inevitable or might be limited through the use of broad or ever-changing proxies. Nonetheless, this discussion emphasizes that transparency could have a substantial cost, lead to the failure of accurate predictors, and thus decrease welfare.

## **Governing Algorithms: The Fairness Argument**

I have argued that transparency in algorithmic sorting does not necessarily enhance welfare. But even if it did, algorithmic sorting could still generate unfairness. The discussion below distinguishes between three categories of unfairness: (a) unfair transfers of wealth; (b) unfair differential treatment of similar individuals; and (c) unfair harms to individual autonomy.

### *Unfair transfers of wealth*

Algorithmic analyses of personal information generate several forms of unacceptable wealth transfers between distinct social groups. One is the transfer of wealth from individuals and consumers to firms—or in our context, from credit recipients to lenders. Another type enables the transfer of wealth among various social groups—particularly from those of lower socioeconomic status to those of a higher socioeconomic status. Finally, it enables transfers that systematically harm minorities and other protected groups. These three issues call for a separate discussion and particular examination of the role automation and opacity play in their overall impact.

The transfer of wealth from consumers to firms (or from credit recipients to lenders) occurs when the social and psychological insights gained by the automated analysis of personal data are abused (see also Calo 2014; Bargill 2012). In the context of credit scoring, the analysis of personal data could allow lending firms to manipulate and lure consumers into seductive lending or detrimental refinancing schemes that are profitable to the financial institution. For example, loan recipients might pay for nonessential services or agree to one-sided, adverse contractual terms. Note that this concern does

not arise from the algorithm's data or processing errors but rather from its ability to understand and predict human conduct, perhaps too well.

Does opacity exacerbate this problem, and would transparency solve it? Oren Bargill explains that enhanced disclosure can educate the public on the unfair nature and extent of the offers made. In addition, data transparency enables competition between firms. In doing so, it could motivate competing firms to offer relevant consumers (who were singled out as prone to manipulation) more attractive options that do not amount to dubious wealth transfers.

Yet transparency, or disclosure-related solutions, might prove insufficient and amount to be a mere political compromise (Teichman 2014). Indeed, the nontransparent nature of the algorithmic processes need not be blamed for generating these forms of unfair outcomes. Other measures might mitigate this concern and should be considered, such as prohibiting the use of aggressive and seductive marketing schemes. In the context of consumer credit rating, limits on the aggressive marketing of problematic financial instruments, such as those including balloon rates, could be implemented. In sum, this "unfairness" issue is serious but might be directly addressed through transparency or other measures.

Transfers of wealth *between consumers* can unfold in various ways. In the context of generating credit scores, let us assume that the partially opaque current framework only allows some knowledgeable segments of the population to understand the workings of the underlying method and therefore "game the system" and access cheaper credit. In this context, the sophistication required pertains to the technical understanding of the system, in addition to broader familiarity with financial and regulatory mechanisms that provide advantages in many other contexts—knowledge that is often associated with financial wealth and high socioeconomic status. In this way, low-cost credit, which turns out to be available to the sophisticated consumer, is at times subsidized by other segments of the population—transferring from the "have-nots" to the "haves." Apparently, the concern arising from this potential wealth transfer is linked to the automated nature of the process. Such processes, as opposed to human discretion-based decisions, are arguably more susceptible to being examined, learned, and gamed.

This phenomenon may be true of any distribution system—automated or manual. It is certainly true of systems that rely upon human discretion. Here, too, high socioeconomic status communities have substantial advantages. In the context of credit, members of stronger groups generally maintain social capital that will facilitate their access to cheaper credit. In other

words, they will “know” someone at the bank or credit agency who will assist them in receiving a better credit score and eventually lower-cost credit. Automated processes in fact *limit* these forms of advantages, particularly the inequality and unfairness gained from possessing “social capital.” Nonetheless, the inequality associated with automated ratings is unique and severe. While “Old World” forms of influence and corruption could be battled using conventional methods of external audits and internal discipline, automated differentiation between individuals, which creates knowledge and sophisticated barriers to its understanding, generates new forms of unfairness.

How do transparency and opacity impact this specific unfairness-based concern? On the one hand, the lack of transparency enhances this concern and exacerbates unfairness, because it allows for the haves to benefit from private information while the have-nots remain clueless. Thus, enhanced transparency could play an important role in mitigating unfairness. Disclosure practices could aid public examination of the credit allocation process, perhaps limiting the ability to game the system in a way that generates inequality and blocks unfair wealth transfers.

Yet, counter to the previous comment, transparency could also exacerbate this unfairness-based concern. This assertion is at first blush confusing, as we often consider transparency (or in metaphoric terms “sunlight”) to be a powerful disinfectant. Yet, transparency works both ways: the public gains more information but, as a result, so do special interest groups. Transparency allows special interest groups to act quickly and influence decisions—actions that often bring about unfair outcomes to weaker population segments. For this reason, budgetary discussions are held in secret and only disclosed after matters are concluded (Vermeule and Garrett 2006).

As an example, consider the prospect of fully transparent (and automated) credit scoring systems. With these in place, special interest groups could quickly move into action and try to influence the process so that specific factors will not be considered a problematic proxy when formulating the credit score (e.g., lobbying by discount store owners to remove purchases at these stores from the list of negative factors). Similarly, groups could lobby to include membership in specific associations as a signal of creditworthiness (consider unions as well as the American Medical or Bar Association lobbying on behalf of their members so that membership in these groups indicate creditworthiness). Lobbying obviously increases unfair outcomes of the processes mentioned because it facilitates a biased decision-making process that systematically benefits stronger and well-organized social segments (and thus is unfair to weaker segments).



While this pro-opacity point might be argued with various degrees of success in almost all contexts involving the planning of public policy, it is worth emphasizing in the context of governing algorithms and automated processes. These processes promise detachment from political and economic tensions and influences. Yet, transparency can potentially undermine the promise of any form of insulation and subject these automated processes to pressures that commonly lead to unfair outcomes.

To summarize, the role and effect of transparency in this context are unclear, and so solutions other than enhanced disclosure must be explored. One solution may be to provide secure government-backed loans and subsidies to individuals with low socioeconomic status in order to compensate them for being excluded from the credit market and to limit the wealth transfer.

A final segment of this discussion on unfairness must address potential discrimination against “protected” groups (such as discrimination premised on race, religion, or gender; Pasquale 2015; Citron and Pasquale 2014). Such forms of inequality are unacceptable for various social reasons (Zarsky 2014). Indeed, here our discussion moves from broader notions of fairness to antidiscrimination theory and policy. It is best to split this argument into three specific claims that merit mention: *First*, in most cases, automated algorithm processes that use race (or other forbidden factors) as factors to decide upon various allocations, are socially unacceptable. *Second*, a skewed and biased data set may cause outcomes of the algorithm process that discriminate against protected groups. This often results from human biases in measurement or other past wrongs that might lead to overrepresentation of some forms of negative data about minorities in the data sets. *Third*, process outcomes might generate a disparate impact (i.e., implicating a racial minority to a greater degree than their representation in the general population)—a process that at times is considered socially unacceptable as well (Citron and Pasquale 2014; Zarsky 2014).

Does the automated nature of the algorithmic practices exacerbate unfairness and discriminatory concerns (for a similar discussion see Zarsky 2012)? All three problematic dynamics may unfold in any decision-making setting, manual or automated. In fact, automation can potentially mitigate concerns of these forms of discrimination, given its ability to partially remove human touch, eliminating hidden or even subconscious biases from the process.

Nonetheless, all three issues generate serious concerns and these problematic practices should be outlawed in many cases (Zarsky 2014). However, each concern mandates a separate response. The first (discriminatory

proxies) calls for absolute prohibition and enforcement. The second (biased data sets) requires that the data sets used include accurate data and are updated frequently. The third (disparate impact) mandates further, intensive research into the nature of the algorithmic process, whether it features a disparate impact, and why.

Beyond these measures, it is important to discuss the impact transparency (or lack thereof) has on these forms of unfairness- (and discrimination-) based concerns. In this context, the answer to this inquiry is relatively simple—the inherent opacity found in the algorithmic processes generates distinct problems related to the fairness issue at hand. As noted, there are at least three distinct causes for this form of discrimination, each provoking a different response. However, opacity impedes inquiry into the nature and reason of discrimination at any specific juncture. The lack of transparency makes it nearly impossible to establish whether a faulty proxy, data set, or another reason has caused discrimination. However, applying transparency measures allows the examination of the nature of the data sets, the factors used, and the final outcomes of the scoring process (which, if found to be controversial, would lead to repeated inquiries of the previous noted factors). With that, internal audits conducted by designated and trusted regulators could most likely sufficiently resolve this set of concerns.<sup>6</sup>

To conclude, algorithmic governing and sorting bring forth a variety of unfairness concerns and dubious wealth transfers. In some limited instances, they are of greater concern and require specific regulatory steps. However, these steps most likely do *not* involve limiting the automated process. Transparency often, but not always, helps to overcome such concerns.

### *Unfairness and treating similar individuals differently*

According to a different fairness-related argument, as a result of the algorithmic process, individuals might be treated differently than their peers—people similar to them in every *relevant aspect*—on the basis of irrelevant differences. For instance, in the credit context, specific purchases or particular behavior could deem an individual a higher risk for default when, in fact, these elements have nothing to do with this particular individual's tendency to repay loans. Or, a specific individual should not have been considered to meet the specific criterion indicating risk given a certain unique circumstance. While this argument appears similar to the efficiency-related concerns noted above, it is distinct in that it pertains to instances in which the overall selection strategy enables efficient and reasonable

identification and mitigation of credit risk as the factors are relevant for *many* individuals. For *some* individuals (yet not enough to render the entire process inefficient), the factors are irrelevant or incorrect. Thus, the process proves to be efficient overall, but in some instances, and for some people, unfair. In this context, it is clear that this problem results, *inter alia*, from automation—leading to the development of models that recommend treating equals differently and renders their inspection and questioning extremely difficult.

Given various mitigating factors, the negative reach of this problem might not be as severe as one would actually believe. For instance, it would cause limited concern in an environment where a variety of algorithmic decisions constantly impact upon our lives in unrelated contexts. In such cases, if the problems do not systematically and disproportionately affect one population more than the other, unfairness in one context will be offset by fairness in other contexts. Furthermore, unfairness in one context might be balanced by other instances in which that individual profits from a beneficial error (allocating goods he or she did not deserve). Therefore, one suggestion to mitigate this unfairness concern is to assure that the various algorithmic systems each individual encounters in a lifetime are distinct from one another. Thus, if an individual is treated unfairly by one, it is possible that he or she will be treated fairly by another. Hence, different individuals will be treated adversely at different times and places—creating an environment where there is a chance that all negative dynamics could be offset. Note that this argument only holds when the various algorithmic interventions are unrelated events and therefore the errors are nonsystematic and do not repeat among platforms. Indeed in many other social setting, this is not the case—it is the same individuals that are discriminated against time and again, with each event reinforcing a subsequent discriminatory trend. Yet the situation discussed here features computer-driven processes that might lead us to believe such reinforcement will not unfold. Such a process limits the role of human discretion that might bring bias with it. In addition, each analysis examines the data sets *de novo*—without necessarily accounting for the distinctions applied in other settings but rather relying on the data itself. Of course, when errors are replicated in various settings this mitigating argument cannot hold.

Finally, let us examine the relationship between this fairness-based problem and transparency. At face value, transparency is essential in constraining this concern. Knowing the factors used in the process, and the reason for their use, can support counterclaims to appeal decisions. Various forms of disclosure will limit this fairness-based concern, yet possibly at the price of

simplifying the automated process and compromising its accuracy. In contrast, transparency compromises the trade secrets of those operating the algorithmic analyses. It is therefore unclear how enhanced disclosure requirements will affect competition in the algorithmic processes market—competition which is of great importance to battle this specific concern; it could enhance it by promoting the flow of information regarding these practices, or stifle it by dampening incentives to innovate, given the prospect of sharing future developments. This is yet an additional matter that must be examined further.

To conclude, this final concern generates a strong argument against algorithmic decision making and calls for greater transparency. With that, it is limited to those instances in which the process relies upon applying inaccurate and irrelevant factors.

### *Unfairness and autonomy-based harms*

Algorithmic decision-making processes raise a crucial additional set of fairness-based concerns. These are autonomy-related concerns that also involve harms to individual dignity. Algorithmic decision-making processes often use personal information without the person's informed consent. Moreover, the analysis also affects individuals in a seemingly arbitrary manner (“punishing people for things they did not do” is a common complaint). It does not enable their understanding of the process's inner workings, or how it impacts on and is impacted upon by their daily lives. In addition, these processes create an environment in which individuals have a limited ability to question the process or submit corrections, and many such concerns have been identified by the law. The EU Data Protection Directive (the “Directive”),<sup>7</sup> for instance, usually provides subjects of ratings with the right to block usage of their personal data without consent (Article 7). Article 12 of the Directive also requires that subjects receive a detailed account of data-based decisions affecting them.

Several of these concerns are linked to the *automated* nature of the process, which could contribute to the sense of arbitrariness it generates. An automated response also limits receipt of a concrete, human explanation of the outcome and any chance of fully understanding the process affecting the individual. Furthermore, almost all of the concerns noted here are linked to the *opaque* nature of the algorithmic process. Lack of transparency leads affected persons to speculate that algorithmic decision making is arbitrary. Opacity further hinders their ability to question the outcome and understand the process. Here too, transparency measures—such as disclosing which

factors were used or the rate of statistical error in predicting the outcome—could be adopted to mitigate these concerns. Yet, to a certain extent, these concerns are inescapable when opting for an (often automated) algorithmic analysis with inherent complexities.

As opposed to the arguments set out above, these issues and concerns cannot be argued away, technically fixed, or corrected by payments and subsidies. They represent real concerns and harms, even though they are somewhat abstract by nature. Therefore, this final set of concerns generates a substantial challenge to the automated and opaque algorithmic process and the policy decisions behind it. With that, the impact of this analytical challenge is not simple or clear.

## **Conclusion**

This article examines some of the arguments often addressed in the public discourse regarding the problems with, and regulation of, algorithmic decision making. The analysis presents progress in proposing various solutions to mitigate substantial concerns. Solutions include abandoning algorithmic decision making entirely, applying forms of transparency, and even imposing indirect remedies such as governmental subsidies and loans, as well as enhancing diversity and competition in relevant settings. This comes in addition to other innovative solutions that have been recently proposed that include introducing audit trails to the algorithmic process or “interactive modeling,” which allows individuals to gain a better understanding on how their actions impact upon the algorithmic response (Citron and Pasquale 2014).

## **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## **Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## **Notes**

1. For a similar analytical move, see Zarsky 2013a.
2. I elaborate on this point elsewhere, see Zarsky 2013a.
3. Such requirements are indeed part of the regulation of the credit ranking process. As Pasquale explains, amendments to the fair credit reporting act (or FCRA,

which governs credit scoring) require credit agencies to provide individuals, upon their request, with the four most dominant factors that led to the relevant decision. This requirement implicitly mandates that those initiating the automated prediction process must be able to explain their actions internally (Pasquale 2015). It is interesting to note that this provision is quite exceptional in US law, although many similar provisions exist in European law. These latter laws often require those initiating an automated computerized process to provide the reasons for the outcomes it generated when an individual is substantially affected. See, for instance, Article 12 of the EU Data Protection Directive.

4. A discussion of costs will include examining the actual out-of-pocket costs the project entails, as well as the other, more distant, costs that might result from it, such as privacy harms. Neglecting to examine this part of the equation does not assume, of course, that such costs are negligible. However, it could assume that comparable costs will unfold in all forms of analysis.
5. This point is addressed in depth elsewhere, see Zarsky 2012.
6. There are distinct reasons to engage in only limited transparency regarding some of these elements and limit disclosure only to a limited circle of regulators and experts. See discussion in Zarsky (2013b).
7. Council Directive 95/46/EC, On the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, 1995 O. J. (L281) 31.

## References

- Bargill, Oren. 2012. *Seduction by Contract: Law, Economics, and Psychology in Consumer Markets*. Oxford, UK: Oxford University Press.
- Becher, Shmuel, and Zarsky Tal. 2014. "Seduction by Disclosure." *Jerusalem Review of Legal Studies* 9 (1): 72-86.
- Calo, Ryan. 2014. "Digital Market Manipulation." *George Washington Law Review* 82 (4): 995.
- Citron, Danielle Keats. 2007. "Technological Due Process." *Washington University Law Review* 85 (6): 1249-313.
- Citron, Danielle Keats, and Frank Pasquale. 2014. "The Scored Society." *Washington Law Review* 89 (1).
- Crawford, Kate, and Jason Schultz. 2014. "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms." *Boston College Law Review* 55 (1): 93.
- Hunt, Robert M. 2005. "A Century of Consumer Credit Reporting in America." FRB Philadelphia Working Paper No. 05-13. Accessed September 24, 2015. <http://ssrn.com/abstract=757929> or <http://dx.doi.org/10.2139/ssrn.757929>.

- Jolls, Chistine, Cass R. Sunstein, and Richard Thaler. 1998. "A Behavioral Approach to Law and Economics." *Stanford Law Review* 50 (5): 1471-550.
- Lenard, Thomas M., and Paul H. Rubin. 2013. *The Big Data Revolution: Privacy Considerations*. Accessed September 24, 2015. [http://www.techpolicyinstitute.org/files/lenard\\_rubin\\_thebigdatarevolutionprivacyconsiderations.pdf](http://www.techpolicyinstitute.org/files/lenard_rubin_thebigdatarevolutionprivacyconsiderations.pdf).
- Meadow, William, and Cass R. Sunstein. 2001. "Statistics, Not Experts." *Duke Law Journal* 51 (2): 629-46. Accessed September 24, 2015. <http://scholarship.law.duke.edu/dlj/vol51/iss2/2>.
- Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA, US: Harvard University Press.
- Perry, Ronen, and Tal Zarsky. 2014. "Queues in Law." *Iowa Law Review* 99 (4): 1595. Accessed September 24, 2015. [http://ilr.law.uiowa.edu/files/ilr.law.uiowa.edu/files/ILR\\_99-4\\_Perry-Zarsky.pdf](http://ilr.law.uiowa.edu/files/ilr.law.uiowa.edu/files/ILR_99-4_Perry-Zarsky.pdf).
- Teichman, Doron. 2014. "Too Little, Too Much, Not Just Right: Seduction by Contract and the Desirable Scope of Contract Regulation." *Jerusalem Review of Legal Studies* 9 (1): 52.
- Vermeule, Adrian, and Elizabeth Garrett. 2006. Transparency in the Budget Process. Accessed September 24, 2015. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=877951](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=877951).
- Zarsky, Tal Z. 2012. "Governmental Data Mining and its Alternatives." *Pennsylvania State Law Review* 116 (2): 285-330.
- Zarsky, Tal Z. 2013a. "Transparent Predictions." *University of Illinois Law Review* 2013 (4): 1503-69.
- Zarsky, Tal Z. 2013b. "Mining the Networked Self." *Jerusalem Review of Legal Studies* 6 (1): 120-136.
- Zarsky, Tal Z. 2014. "Understanding Discrimination in the Scored Society." *Washington Law Review* 89 (4): 1375.

### Author Biography

**Tal Zarsky** is a professor at the University of Haifa – Faculty of Law. His research focuses on Information Privacy, Cyber-Security, Internet Policy, Social Networks, Telecommunications Law and Online Commerce, He also teaches and studies Contract and Property law and theory.