

STRATEGIC GAMES AND ALGORITHMIC SECRECY

*Ignacio N. Cofone and Katherine J. Strandburg**

We challenge a claim commonly made by industry and government representatives and echoed by legal scholarship: that algorithmic decision-making processes are better kept opaque or secret because otherwise decision subjects will “game the system”, leading to inaccurate or unfair results. We show that the range of situations in which people are able to game decision-making algorithms is narrow, even when there is substantial disclosure. We then analyze how to identify when gaming is possible in light of (i) how tightly the decision-making proxies are tied to the factors that would ideally determine the outcome, (ii) how easily those proxies can be altered by decision subjects, and (iii) whether such strategic alterations ultimately lead to mistaken decisions. Based on this analysis, we argue that blanket claims that disclosure will lead to gaming are over-blown and that it will often be possible to construct socially beneficial disclosure regimes.

Nous contestons une affirmation couramment véhiculée par certains représentants de l'industrie et des gouvernements et qui est parfois reprise dans la littérature juridique, soit l'idée selon laquelle il est préférable de garder les processus de décision algorithmiques opaques ou secrets, faute de quoi les sujets des décisions pourraient « déjouer le système », ce qui conduirait à des résultats inexacts ou injustes. Nous montrons que l'éventail des situations dans lesquelles les gens sont capables de se servir des algorithmes de prise de décision pour son propre avantage est étroit, même lorsqu'une quantité substantielle d'information a été divulguée. Nous analysons ensuite comment discerner les situations où il est possible de détourner les algorithmes en fonction (i) du degré de connexion entre les indicateurs de décision et les facteurs qui, idéalement, détermineraient le résultat, (ii) de la facilité avec laquelle ces indicateurs peuvent être modifiés par les sujets des décisions et (iii) de la possibilité que ces modifications stratégiques conduisent finalement à des décisions erronées. Sur la base de cette analyse, nous soutenons que les allégations générales selon lesquelles la divulgation conduirait à des détournements sont exagérées et qu'il sera souvent possible de mettre en place des régimes de divulgation socialement bénéfiques.

* Ignacio N. Cofone, Assistant Professor, McGill University Faculty of Law. Katherine J. Strandburg, Alfred Engelberg Professor of Law, NYU School of Law. We thank Sebastian Benthall, Cassi Carley, Alessa Dassios, Guiseppe Dari-Mattacci, Ashley Gorham, Yafit Lev-Aretz, John Nay, Garry Gabison, Mason Marks, Julia Powles, Ira Rubenstein, Madelyn Sanfilippo, Mark Verstraete and Ari Waldman for their helpful comments. The paper also benefited from numerous comments on presentations at the NYU Privacy Research Group, NYU Information Law Institute, and the Programming Governance & Governing Programming Conference at McGill University. Katherine Strandburg acknowledges the generous support of the Filomen D'Agostino and Max E. Greenberg Research Fund. Ignacio Cofone is grateful for the support of the Social Sciences and Humanities Research Council (Insight Development Grant) and the Canadian Institute for the Administration of Justice Charles D. Gonthier Research Fellowship. We also thank Malaya Powers for her outstanding research assistance.