Some new approaches for successful use of data analytics in the public sector

Paper presented to the Northeast Conference on Public Administration

November 2020

Authors

Professor Timothy O'Loughlin, Carnegie Mellon University Australia

Timothy O'Loughlin is a Professor of the Practice of Public Policy and Management at Carnegie Mellon University Australia. Tim returned to university work in 2006. His immediate positions before joining CMU-A were running agencies in the South Australian Government where he specialized in infrastructure, transport, climate change and renewable energy policy.

Rachel Bukowitz, Carnegie Mellon University, Heinz College for Information Systems and Public Policy

Rachel Bukowitz is completing her postgraduate studies in Public Policy and Management at Carnegie Mellon University. Her studies are focused on leveraging data and technology to analyze and implement policy solutions. Her interest is concentrated in evaluating Big Data applications in governments.

ACKNOWLEDGEMENTS

We would like to thank Dr. Emil Bolongaita, Head of CMU-A and Distinguished Service Professor of Public Policy and Management, and Dr. Danura Miriyagalla, Deputy Head and Professor of Practice, CMU-A for their comments on this paper.

Abstract

Governments using data analytics will find themselves increasingly drawn into creating social licence for these applications. This paper develops some strategies for achieving it. The need will be greatest where two conditions are present. First where data analytics are used to predict rather than describe or prescribe. Second, where such prediction is used by governments when exercising their coercive powers. The paper considers two such cases briefly – predicting children at risk of neglect and abuse; and predicting recidivism risk in the criminal justice system. Each has drawn significant criticism, helping to precipitate discussion around the requirements for social licence. The paper contends that conventional techniques for creating

licence are unlikely to prove effective due to the high degrees of opacity and complexity of the tools allied with unclear accountabilities for establishing the social values embedded in them. However, the need to infuse the application of these techniques with democratic principles is unavoidable for social licence to be established. The conclusion is that unconventional democratic forms, including deliberative and direct democracy, are likely to prove more successful than representative democracy in establishing that licence and thereby realizing more fully the potential contribution of data analytics to better government.

Introduction

The problems being experienced by corporations in establishing social license for their data analytics applications are well-known. They have spawned calls for greater regulation, including calls from within the industry itself with Google calling for “common rules of the road” for regulating (Kharpal 2020); Zuckerberg arguing for outright regulation of the internet (Zuckerberg 2020) and Amazon, IBM and Microsoft refusing to make facial recognition software available to police departments until governments put in place legislation to govern the use of the technology (Vanian 2020). Such calls can be seen as the private sector looking to draw upon citizens’ trust in government to help make up for the lack of trust in its own activities and the consequent problems created by that for their social licence to operate.

While governments may be better placed than private companies to create trust, use of data analytics in the public sector presents governments with deeper problems because of differences in the relationships between people and companies and people and their governments. For most, the customer relationship is the extent of their relationship with companies. The relationship between people and their governments is more complex as people relate to government as clients, citizens, customers and subjects (Mintzberg 1996). When it comes to the use of data analytics, the last of these roles is the most problematic as it is here where government legitimacy, justice and fairness face their greatest tests.

There are many examples of data analytics making for better government, from broad capability improvements in government efficiency (McKinsey 2011); satisfaction with government services (Mehr 2017); and in enabling eParticipation (Sæbø, Rose, & Skiftenes 2008) through to specific innovations such as helping make visible the ‘credit invisible’ (Executive Office of the President 2016); and California retrospectively expunging marijuana convictions following law reform (Jackson, A 2019).

Examples of failure associated with use of data analytics by governments abound too, with those being used to support coercive powers over-represented - Ireland’s cervical cancer screening scandal (Scally 2018); the UK Government’s Nudge Unit tool for evaluating schools (Williamson 2018); Illinois’s abandonment of the use of a predictive tool for child protection; (Jackson and Marx 2017); the Australian Government’s twice-abandoned Robodebt welfare payment recovery scheme (McPhee 2017); and the still unresolved issues associated with genomic testing of children (Wilfond 2012).]

A good illustration of the controversy that can be associated with the use of data analytics by governments is the demise of the UK Government’s effort to create a centralised medical information repository. The scheme had to be abandoned in 2016, four years after its enabling legislation had been passed and following two years of controversy. It is argued that the scheme’s failure was attributable to a community trust deficient; its potential disruption to traditional doctor-patient relationships; and failure to establish it as an unambiguous public good (Carter, Laurie & Dixon-Woods 2015). In short, a lack of social licence.

No single case is the same but the failures share several common elements. Within government, defective design of tools as well as poor management of the data and its mining are common elements. The impact is exacerbated by communication failures between IT designers and their political overseers whose interest in, or ability to grasp, the technical issues can be limited. For instance, McPhee (2017) has argued that the massive increase in risks from a fifty-fold escalation of the rate of data matching by the Australian Government's welfare agency for its Robodebt debt recovery project was unlikely to receive sufficient attention by decision-makers given the political imperative to implement it.

There are also significant exogenous factors working against successful application of data analytics in public sector settings. These include falling trust in governments, greater access for citizens to information, greater demands for participation, and diminished standing of "experts", brought on partly by the increased citizen participation in policy-making (DeMarchi, Lucertini & Tsoukias 2016, McKee 2010, Price 2000, Tsoukias 2013).

The use of data analytics provides a tougher test for establishing social license than other applications, most of which are descriptive analytics used to support service delivery. In some cases, failing to create social licence is the final product of the shortcomings in technical performance as was the case with Ireland's cervical cancer tool and Australia's Robodebt. For others, such as the UK's care.data, the technical performance may have been sound but the strategies for creating social licence were flawed or non-existent. In all cases, the criticality for success was underappreciated and the difficulty in creating it underestimated.

This paper begins by looking at social licence as a concept. This inevitably raises the problem of the opacity of the tools as the greatest impediment to achieving social licence and the conventional wisdom that the remedy is greater transparency. The use of the concept of transparency in this context is found wanting, suggesting a need to look at different approaches. Some ideas for these alternative approaches are presented in the last section of the paper.

Concepts: social licence and transparency

The original concept of social licence emerged from the mining industry and remains an important tool for constructing the legitimacy of mineral extraction (Parsons, Lacey & Moffat 2014). Its use has mushroomed within the private sector more generally and is now a mainstream concept. According to Gehman (2017) media use of the term in articles grew from 10 mentions per year up to 2002 to more than 2,000 in 2016.

There are multiple definitions, two of which suit the purposes here. According to Gunningham, Kagan and Thornton (2004) social licence constrains executives "to meet the expectations of society and to avoid activities that societies… deem unacceptable." Another from the Australian Centre of Corporate Responsibility sees social licence in terms of "the level of acceptance or approval continually granted to an organization's operations or project by local community and other stakeholders". Both of these definitions are helpful because they can be applied to the activities of governments as well as the private sector.

The use of the social licence test is less common for governments. After all, governments can create their own legal licence so why the need for social licence as well? The reality is that the disadvantages of relying on legislative authority are becoming more pronounced. As the speed of innovation in delivery of government services by both the private and public sectors increases, legislation becomes a more inflexible tool, unable to be adjusted quickly or prospectively to create the necessary authority for the potential of that innovation to be realised. It is instructive that the legislation for the aborted UK health data project was passed two years ahead of the attempted introduction of the project yet demonstrably failed to create the necessary authority within the general public and among medical practitioners for it to succeed (Carter, Laurie & Dixon-Woods 2015).

The core need for social licence is explained neatly by Brauneis and Goodman (2017):

> An individual can be denied parole or credit, fired, or not hired for reasons that she will never know and which cannot be articulated. In the public sector, the opacity of algorithmic decision making is particularly problematic, both because governmental decisions may be especially weighty and because democratically elected governments have special duties of accountability. (103)

If opacity is the problem, the intuitive antidote is transparency and the calls from that come regularly from the media, politicians, and scholars (Brill 2015, Dwork et. al. 2011, Carlson 2017). Braunies and Goodman argue that ‘It will be possible to assess a predictive algorithm's politics, performance, fairness, and relationship to governance only with significant transparency about how the algorithm works’ (128). For them, the test is not absolute transparency but whether citizens have transparency ‘sufficient to approve or disapprove of the algorithm's performance’. (132)

However, the reality is that even this lesser test is virtually unachievable:

> Based on the evidence we have received, we believe that achieving full technical transparency is difficult, and possibly even impossible, for certain kinds of AI systems in use today, and would in any case not be appropriate or helpful in many cases. (House of Lords 2018:38)

Finding a way through this requires placing the pursuit of transparency into some context. It has been described as Big Data’s Achilles heel as it is a necessary condition for results to be confirmed or refuted (Cohen 2013). However, the successful application of data analytics in public sector contexts is contingent on social licence which, in turn, is contingent on justice and fairness. Other criteria, such as the accuracy of the analytic tools and establishing the credentials of the tools as public goods, serve along with fairness and justice as instrumental means for achieving social licence. Transparency is another such criterion as it can be seen as a contributor to justice and fairness. It also serves other values such as representation (Robbins and Henschke 2017) and as “...a necessary step to accountability” (Eaglin 2017:111).

The difficulty with transparency arises when its instrumental role is overlooked in favour of transparency being seen as the only strategy for achieving social licence or even as an end in itself. The problem can be illustrated by considering two cases of the use of predictive tools in public sector contexts.

The first case is the application of a predictive tool in Allegheny County Pennsylvania to assist social workers making assessments of children at risk of abuse and neglect. The brief history begins in New Zealand in 2015 when officials, attempting to use a like tool, proposed an observational study to test the tools predictive algorithm's accuracy on 60,000 at risk children over a two year period. The then Minister famously commented 'Not on my watch! These are children not lab rats' and that she 'could not fathom what her officials were thinking' (Stuff 2015).

A possible explanation might be found in a similar outcome reported for the use of an algorithm in Boston to better match transportation schedules with school starting times. The application of the tool would have seen some children needing to start school at 7.15 a.m. Masuga (2018) comments that "policy makers and administrators arguably lost sight of constituents' concerns in pursuit of algorithmic efficiency". It is a perceptive insight, suggesting administrators confused the means with the ends. More importantly, it also suggests a tendency among both administrators and developers to seek social licence for their predictive tools in terms of the tool's utility but as Moore has pointed out "...once a public agency engages state authority, justice and fairness in both operations and outcomes become just as important as efficiency and effectiveness…" (2016 P.55).

Much the same methodology was used with greater success in the U.S. by Allegheny County, Pennsylvania. Here the proponents set up the process for creating social licence more systematically using a 12-step process. Community engagement began as the third step straight after development had begun. This was followed immediately by ethical review carried out before the project was launched and case workers trained in its use. The practical reality of the fact of the opacity of this and other tools was recognised by the researchers:

> In addition to knowing what the outcome is going to be, one may also need to know why this is the case. Even the simplest prediction models can only speak to the question of why in a limited sense. (Chouldechova et.al. 2018:55)

Nonetheless, the researchers considered that transparency was achieved:

> Throughout the project, the research team and County leaders had a strong commitment to transparency. They met with community groups, stakeholders and families who were in the welfare system multiple times. This has led to a strong community acceptance of the project. (13)

The problem of conflating transparency as an end value with its instrumental value has been noted already. To this can be added the problem of ambiguous use of the concept.

What is meant by the "strong commitment to transparency"? If it is not possible to explain why it came to the conclusion it did, what part of it is transparent? A second conceptualisation implied in the above is that it refers to transparency about the development process and planning for use of the tool rather than transparency about the tool itself. It seems to be this conceptualisation that Keddell (2019) had in mind when observing that "the Allegheny tool has had much more transparency than others and considerable community involvement, a public technical report, and ethical and impact evaluations" (4).

A subsequent study by a related team (Brown et.al. 2019) reported on workshops conducted with 83 participants drawn from affected communities with the aim of 'designing a blueprint to aid government agencies in improving community comfort levels with algorithmic decision-making' (2). The study found that opacity was a serious issue for the participants who asked specific questions about the actual operation of the tool but the authors cautioned that 'simple disclosure of this information is unlikely to have much effect on promoting relationship building' (10).

The problem here is that strategies for achieving community 'comfort' and forming 'supportive relationships' which engender trust in the fairness and justice of the use of the tool are being elevated above peoples' legitimate democratic rights.The purpose of overcoming opacity here is to help build faith in the operation of the tool. This purpose is elevated above the more fundamental purpose of serving peoples' democratic rights as they are not being provided with 'sufficient to approve or disapprove of the algorithm's performance'. Overcoming opacity is defined as meaning getting people to trust in the tool just as they trust getting on a plane without knowing how it works. That may be laudable but it is no reason to not disclose information about how the tool works, particularly to those subjected to the exercise of coercive powers that the tool helps to inform.

A third conceptualisation might be that transparency is achieved when the normative considerations implicit in the construction of the tool are made explicit and resolved by communities rather than developers. Both the process and values conceptualisations of transparency seem to fall well short of the aforementioned test set by Braunies and Goodman. At a minimum, this would require knowing the programming rules of operation embedded in the algorithm originally as well as how the algorithm validates the invention of any subject rules created by machine learning.

Even if these less demanding tests of transparency are accepted, there seems to be real difficulties in achieving them in practice. While the limited test of process transparency may be satisfied for the community, the groups that most need to be satisfied are those using it and those subjected to it. Eubanks (2018) found considerable shortcomings with the Allegheny algorithm in this regard.

A fourth conceptualisation of transparency for tools used to support coercive powers arises in the second case which concerns the use of predictive tools to assist judges in determining sentences. Specifically, the tools provide data-based assessments of the likelihood of re-offending. They are more commonly used for assessing bail applications, though their use for augmenting considerations to be taken into account for sentencing continues to grow (Kehl et. al. 2017).

One such instance occurred with the use of a tool – the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) algorithm – in preparing a pre-sentencing investigation (PSI) report for the sentencing of Eric Loomis, a man who pled guilty in Wisconsin to two charges related to a drive-by shooting case. Loomis' PSI specifically included his score from the COMPAS algorithm. On the day of Loomis' sentencing, the court stated:

> You're identified, through the COMPAS assessment, as an individual who is at high risk to the community. In terms of weighting the various factors, I'm ruling out probation because of the seriousness of the crime and because your history, your history on supervision, and the risk assessment tools that have been utilized, suggest that you're extremely high risk to re-offend. (State v. Loomis, 2016 P.8)

Loomis appealed his sentence on process grounds, arguing that the use of COMPAS violated his right to have his sentence based upon accurate information, in part because he was unable to verify the accuracy of his score due to the proprietary nature of COMPAS preventing him from assessing it; the tool's use of aggregated data violated his right to an individualized sentence; and that it improperly used gendered assessments in sentencing (Freeman 2016).

The Wisconsin Supreme Court overruled Loomis's process case arguing that although he "cannot review and challenge how the COMPAS calculates risk, he can at least review and challenge the resulting risk scores set forth in the report attached to the PSI" (State v. Loomis 2016 p.22). The courts furthermore referenced Northpointe's 2015 Practitioner's Guide to COMPAS which showed that the risks scores are mostly based on static factors which Loomis could in fact verify.

The court's referral to the Practitioner's Guide addressed the question of accuracy surrounding COMPAS, but failed to address the fairness (or lack thereof) of using such a tool for sentencing. In its final ruling the court mandated that any forthcoming PSIs containing COMPAS risk assessment scores were required to include a written advisement about the proprietary nature of COMPAS and other of its limitations. Finally, the judges' provided an assurance that COMPAS does not violate a defendant's due process rights:

> 'Ultimately, we conclude that if used properly, observing the limitations and cautions set forth herein, a circuit court's consideration of a COMPAS risk assessment at sentencing does not violate a defendant's right to due process.' (State v. Loomis, 2016 paragraph 8).

The implication is that the community should trust the judge to use the tool wisely in lieu of having the ability to understand how it works:

> …. the Court complemented its principal argument by adding that, even if Loomis' objections were partially right, the judges' expertise would compensate for the inaccuracy of the tools and that the obstacles to relying on the developers' capacity and/or willingness to ensure it might be a serious issue. (Beriain 2018 p.47)

The implication therefore is that one way of satisfying demands for transparency is to substitute it for trusting in the assessments made by others.

An important contribution to this discussion comes from Cynthia Rudin who has argued that 'high stakes decisions' should be made using interpretable models rather than black box models (2019). She finds 'there cannot be an all purpose definition' of interpretability as it is domain specific (P.206) and that 'there is a formidable computational hurdle in designing interpretable models' (P.211). Separately, she and others reverse engineer the COMPAS algorithm to establish that it can be made transparent and that 'the focus on the question of fairness is misplaced, as these algorithms fail to meet a more important and yet readily attainable goal: transparency' (Rudin, Wang and Coker 2019. P.1)

The argument for interpretability takes people closer to having their democratic rights met. However, as Rudin recognises, there is a 'chasm' between interpretability and explainability. Therefore interpretability too is insufficient to provide citizens with 'sufficient to approve or disapprove of the algorithm's performance'. Rudin et. al. go on to argue the case for privileging transparency over fairness 'because there is no clear definition of fairness and competing definitions are largely incompatible' (P.1)

There are, of course, many philosophical definitions of fairness with the most commonly used being Rawls' equating fairness with justice. However, for these purposes, it may be more helpful to consider fairness in instrumental terms. Using Moore's prescription, it may be taken that fairness is achieved when the use of a tool meets the tests of the political authorising environment. This assessment can be applied to 'high stakes decision-making', taken for current purposes as decisions made in the public sector to apply legitimate, coercive powers based on predictive risk modelling such as that used to inform decisions on child protection, bail and sentencing.

The conclusion here is that it is not fairness which struggles with multiple definitions but transparency which suffers from multiple meanings even when used for a single tool and that substituting interpretability for transparency falls short of allowing people to exercise their democratic rights. For these reasons, using transparency as a strategy for establishing social licence seems a dead end. It appears to be unachievable in its most exacting meaning leaving all other tests of it as default options for not achieving perfect transparency. These other tests are transparency about the process being used; transparency about the values embedded in the tools; transparency provided through validation by third parties; and transparency as interpretability all fall short of what is required. Their meaning is ambiguous and the assessments of whether the standards have been met are subjective judgements. The greatest flaw common to all of them is that they are used as vehicles for circumventing peoples' legitimate expectations of fairness and justice by providing them with the wherewithal to inform themselves about the operation of tools used for purposes as serious as denying them parole; sending them to jail; or having their children removed.

Meeting this test is likely to prove unachievable. To take the criminal justice example, Eaglin invites consideration of how releasing the data used in the development of the tools would benefit anyone other than those with data analytic skills, much less judges and defendants - 'Such a requirement could have the opposite result of obscuring tool construction even while making it technically transparent' (p.106)

Achieving social licence

Strategies for achieving social licence need to respond not only to the reality of constrained transparency but also to some other realities.

First, achieving social license is not a binary variable. Unlike the U.K.'s care.data, both cases used here show that data analytics tools can continue to operate with deficits in social license. However, the fact that such applications are being used widely is not to say that their use is uncontroversial or debate over their uses is settled.

Second, the use of data analytics for prediction is more fraught than using it for description. For description, the paramount requirement is for the data to be accurate. The same standards cannot be applied to prediction. Here the requirement is to make it as accurate as possible recognising that this is a matter of degree. By way of example, the landmark study by Kleinberg et. al. (2017) of an algorithmic approach to bail decisions claimed much greater accuracy but the algorithm still only had an Area Under the Curve (AUC) of just 0.707.

Third, a higher standard of social license is required for cases in which an algorithm-generated prediction can be used as justification for the public sector to exercise high levels of coercive powers, such as removing children from their homes or making criminal sentencing decisions. In these cases where the predictions are necessarily less than 100 percent accurate and the consequences of the predictions can have profound impacts on their subjects, achieving social license demands methods of assurance that satisfies not just the public, but particularly users and those subjected to it, that these tools are being deployed in a fair and principled way.

Democracy

Achieving social license for government applications of data analytics generally and PRM specifically is difficult. The inherent intrusiveness of many big data applications; privacy concerns often emanating from spillover effects of private sector transgressions; doubts about the extent of public good attributable in part to their novelty; and a trust deficit for public sector innovation generally all contribute to the challenge. The problem of opacity is perhaps the greatest impediment as it makes ideal solution of genuine transparency practically unachievable.

The opacity problem will only become greater as the tools become more sophisticated. Trade-offs between the predictive performance of tools and their opacity are reflective of a deeper tension between better and more efficient public sector service delivery and democratic accountability. For instance, Chouldechova (2018) reported some success in refining the predictive fairness of the Allegheny tool but observed that this came at a cost of making it more difficult to interpret. (P.10).

This is not to argue against efforts to make tools more transparent. There are plenty of suggestions in the literature. For example, proposals by Cohen (2013) include holding engineers directly accountable to the community and regulators resourcing citizens to run their own interrogations. Like Eaglin, Citron (2007) places heavy emphasis on empowering publics to have a deliberative role in such participation and on the need for audit trails for decision-making which include identifying those responsible for making the decisions. Leonard (2018) wants citizens to have a right to “reasonable inferences” as well as a right for humans to intervene directly. Various contributors in Helbing, Helbing and Caron (2019) propose regulatory interventions based on the state guaranteeing “informational self-determination”; using scientific institutions to act as custodians of data and “algorithms that evade democratic control”; community-owned data banks; and use of referenda for data processing and management issues.

The common thread in these and other proposals is the emphasis on people and communities engaging directly with developers and governments in the social licensing process. This brings with it an implied or stated lack of faith in the ability of representative government to deliver social license. According to Helbing “centralized, top-down control is a solution of the past, which is only suitable for systems of low complexity. Therefore, federal systems and majority decisions are the solutions of the present” (2019 P.80).

The emphasis on direct, participatory democracy in the literature can be seen as symptomatic of declining trust in government generally as well as a lack of regard for the ability of elected representatives to come to terms with contemporary, arcane issues that move too fast for the regulators to respond before the damage is done. Developments in constructing algorithmic tools happen too fast to fit in with electoral cycles leaving policy interventions short of electoral mandates. The design of the institutional forms needed to manage these issues also requires a degree of flexibility not well suited to traditional government. The breadth of participation required for genuine consultation on applications with such ubiquitous impacts can also prove difficult for governments used to dealing primarily with formally-organised interest groups. In short, social licence cannot be achieved without the active involvement of those affected by them.

This is not to argue against any roles of representative government. The democratic authority it can lend to creating social licence cannot be replicated by direct, participatory democracy. The question then becomes what institutional forms might be best suited to delivering social licence? The consensus in the literature coalesces around people and communities involving themselves directly. This may take a number of forms such as plebiscites, co-production, citizen's juries and local committees. Such forms are particularly valuable in those areas of government service delivery, such as public health, where there is a 'democratic deficit' because decision-makers are not elected (Lenaghan 1999).

One form which takes advantage of both the democractic authority of representative government and participatory democracy is the Empowered Participatory Governance model developed by Fung and Wright (2003). It sees expert opinion being made available directly to citizens through a decentralisation process which is not autonomous but co-ordinated from the centre. Krantz sees this taking the form of 'technodemocracy' in which 'citizens contest the dominance of experts and technical staff learn new ways of communicating with the public' (Krantz 2003: 232).

All of these approaches have clear limitations. The first is constraints on the abilities of individuals to participate. The reality is that, no matter how skilled the interlocutors might be, it is not possible to participate in meaningful dialogue without some capabilities and knowledge. People have innate conceptions of justice and fairness but applying them in a consultation on the design of a data analytics tool demands more of them than intuition. They need to know what the technology is capable of; the various options for how it might be used; the immutable legal, ethical and accountability constraints; how results might be evaluated; and the government decision-making processes into which the results of the consultations are to be injected. Most of all, can citizens participating in such a process be relied upon to meet Kant's test of placing the right over the good?

The second, related constraint is the information asymmetry created by technologies generating more information for governments about citizens without a commensurate increase in citizens' knowledge of the state. Robbins and Henschke (2017) argue that this puts a premium on the value of representation but that assumes that the same constraints do not apply to representatives.

A third constraint is the ability of the state to make available effective channels for that consultation. Plainly the use of mass forms of consultation such as plebiscites are unworkable for community engagement processes that need to be highly specific and detailed in their responses. These processes also need to be replicable as a unique one will be needed for the development of each tool.

Against this background, new proposals for creating social license continue to be hatched. Medhora and Owen (2020) have proposed a new international framework for digital governance to mirror the Bretton Woods institutions. Robbins and Henshcke (2017) argue for practice of Value Sensitive Design aimed at embedding specific moral values such as representation into the technologies as the means for ensuring they reflect the will of citizens. Masuga (2018) has observed the developers themselves have worked on "technical recommendations aimed at creating baselines for accountability (P.106). Janssen and Kuk (2016) propose democratizing algorithms by engaging users "...to co-create alongside with expert practitioners in areas of design, software development and machine learning" (P.376). While drawing on concepts of schools as free, civic spaces, Boyte (2017) has raised the concept of using educational sites as a medium for citizen involvement. At the municipal government level, The University of

Pittsburgh has created a task force drawing members from academe, government and the community to create practical guides and best practices for the use of algorithms (Taylor 2020).

These and many other ideas are generally aimed at greater citizen engagement with, and ownership of, the information technologies used to govern them. The problem is that the proposals generally do not come to grips with the constraints referred to in the foregoing of limited citizen capacities and knowledge; inevitable information asymmetries; and limitations on the abilities of governments to create mechanisms that stimulate, collect and use that input in a deliberative way.

A more promising approach is an adaptation of the model developed by Dryzek and Stevenson (2011) for climate change governance. Starting out with a contrast between consensual and adversarial democracies, they find that the environmental performance of consensual states is generally better but that "the close proximity between the state and civil society tends to thwart radical critique, which is necessary insofar as their environmental performance remains inadequate" (P.1866). The authors illustrate the point by contrasting outcomes in two consensual states - Norway and (West) Germany. Norway's public funding of environmental groups had the perverse effect of keeping their memberships low and their critiques of policy moderate. Germany did better out of having more independent groups providing a more strident critique of the political economy of environmental policy.

The conclusion is that poor performance in the ecological area by consensual states makes radical critique a "necessity" and "if consensual states cannot generate this kind of critique themselves, they must import it from elsewhere" (p.1866). So the requirement for effective governance is for the state to have a deliberative public space where ideas are contested. This public space needs to be held at sufficient distance to allow the public space to freely test ideas. The empowered space is where decisions are made e.g. legislatures while the public space is where ideas get tested and debated.

The attraction of such an approach for public sector data analytics is the strong interface between the technical and the normative. In the public space, experts participate with informed and engaged citizens in debating both technical and normative issues . This is no ivory tower as the public space is connected to the empowered space, albeit at a distance.

As an example of how it might work, Drysek and Stevenson cite the meetings of the UN Framework Convention for Climate Change where delegations of national government negotiating conventions for limiting global emissions occurs contemporaneously with a plethora of side events in which academics, members of activist groups, technical experts, bloggers and concerned citizens try out new ideas and have them debated. These debates are at some length from the policy haggling in the main rooms allowing them a freer hand. However, they remain connected to the empowered space, both physically and through pre-existing relationships between activists and officials.

Another example of an empowered space is the provincial-level Maternal Mortality Review Commissions (MMRC) in the United States. MMRC memberships originally consisted of mostly obstetricians, but today have expanded to include a range of representatives from medical and health fields, as well as patient advocacy groups and community organizations (Review to Action). For example, Washington D.C.'s MMRC membership is mostly composed of health and non-governmental representatives, as well as five members from government agencies (D.C. Law 22-111. Maternal Mortality Review Committee Establishment Act of 2018). The D.C. MMRC is structured in a way in which:

> detailed meeting minutes and recordings of the public portion of the meeting will be provided in accordance with the Open Meetings Act. The Committee welcomes public comments and questions and reserves time at the start of each meeting for public engagement. (Open DC Maternal Mortality Review Committee 2019).

D.C.'s mixture of government officials, relevant professionals, and interested individuals from the public allows for free debate and exchanges of ideas across a diversity of backgrounds. Although the MMRCs across the U.S. are structured differently, they are united by their common objective of reviewing the pregnancy-related death cases in their respective states or jurisdictions, and then offering new ideas, recommendations, and reports to the state for reducing maternal mortality rates. The public sphere is still connected to the empowered space as MMRCs' data is shared with the U.S. 's leading public health institute, the Center for Disease Control and Prevention (CDC), as well as state health departments.

Unlike the UNFCCC events which operate at a supra-national level, MMRCs operate at the provincial level, allowing them to engage locally while still maintaining the opportunity to influence national policies. The proximity of the MMRCs to the affected populations (i.e. people impacted by pregnancy-related deaths) gives their data authority while the inclusion of a variety of stakeholders beyond solely government officials and designated experts creates space for honest critique which in turn informs changes in state and national action on maternal mortality.

What form might this take in the process for constructing social licence for data analytics?

That form will certainly not be the hackneyed, undemocratic idea of independent, expert bodies acting as overseers of government intentions. As Eaglin and others have pointed out, the debate is not about the efficacy of these tools. It is about the contest of values, those embedded in the tools themselves and those loaded into the manner of their use. Experts in data analytics are no better placed than any other individuals to participate in this determination.

The catalyst for creating the required public space might be the creation of Bretton Woods-type institutions for data analytics envisaged by Medhora and Owen (2020). The university sector can be expected to spawn activist groups such NYU's Now Center and the University of Melbourne's Centre for Artificial Intelligence and Digital Ethics. However, this organic growth may not proceed fast or fully enough to make the same sort of contribution to data analytics use as the UNFCCC has made to climate change.

More likely, it will come from local groups such as the Maternity Review Commission. As the work of Eaglin, Eubanks, Keddell and others has shown, the impacts of data analytics tools are felt most powerfully by those subjected to them and those doing the subjecting. The normative issues that emerge are likely to be highly specific to the needs of both and need to be resolved in a manner that can be supported by the communities in which they reside.

At any level, it is in the interests of government to stimulate the emergence of such groups. The fundamental requirement for creating contested public space is for government to give them a connection with policy-making. This does not mean a formal, consultative framework but an informal relationship as government recognising that it is in its best interests to provide formative materials such as draft tender documents, development designs and subsequently source codes and so on. It may inflict some damage in the short-term but will help avoid terminal outcomes such as care.data in the medium term.

Conclusion

The empowered space/public space dichotomy seems well-suited to the creation of social license for public sector data analytics. It leaves the elected government in charge, drawing upon its democratic authority while taking advantage of opening of another element of democracy of using debate to persuade and being open to be persuaded. In doing this it draws upon activists with the knowledge and commitment of data analytics techniques to properly hold representative government to account.

Reading list

Angwin J, Larson, J Mattu, S & Kirchner L 2016 'Machine Bias'

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Australian Centre of Corp[orate Social responsibility quoted in Moore, M. "The Question of Social Licence and Regulatory Responsibility." *The School of Public Policy Publications (SPPP)*, vol. 9, 2016.doi:http://dx.doi.org.proxy.library.cmu.edu/10.11575/sppp.v9i0.42574

Babič F and Zuskáčová K."Descriptive and predictive mining on road accidents data," 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Herlany, 2016, pp. 87-92.

Berian, M (2018). Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling. Law, *Probability & Risk*, 17(1), 45–53. https://doi-org.proxy.library.cmu.edu/10.1093/lpr/mgy001

Blacklaws C 2018 'Algorithms: transparency and accountability' *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* http://doi.org/10.1098/rsta.2017.0351

Boyte, H. "John Dewey and Citizen Politics: How Democracy Can Survive Artificial Intelligence and the Credo of Efficiency." *Education and Culture* 33.2 (2017): 13-47.

Brauneis, R., & Goodman, E. P. (2018). Algorithmic Transparency for the Smart City. *Yale Journal of Law & Technology*, 20(1), 103+.Retrieved from https://link-gale-com.proxy.library.cmu.edu/apps/doc/A544510684/ITOF?u=cmu_main&sid=ITOF&xid=d3bd182a

Brill, J 2015 Scalable Approaches to Transparency and Accountability in Decision Making Algorithms Remarks at the NYU Conference of Algorithms and Accountability. https://www.ftc.gov/public-statements/2015/02/scalable-approaches-transparency-accountability-decisionmaking-algorithms

Brown, A., Chouldechova, A., Putnam-Hornstein, E., Tobin, A. and Vaithianathan, R., 2019, May. Toward algorithmic accountability in public services: A qualitative study of affected community perspectives on algorithmic decision-making in child welfare services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

Carlson, A. M. (2017). The Need for Transparency in the Age of Predictive Sentencing Algorithms. Iowa Law Review, 103(1), 303+. Retrieved from https://link-gale-com.proxy.library.cmu.edu/apps/doc/A518901209/AONE?u=cmu_main&sid=AONE&xid=8f40239b

Carter, P., Laurie, G.T. and Dixon-Woods, M., 2015. The social licence for research: why care. data ran into trouble. *Journal of medical ethics*, *41*(5), pp.404-409.

Chouldechova, A., Benavides-Prado, D., Fialko, O. and Vaithianathan, R., 2018, January. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In Conference on Fairness, Accountability and Transparency (pp. 134-148).

Citron D 2007 ‘Technological Due Process’ Washington University Law Review Vol.85, pp.1249-1313
https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview

Cohen, J 2013 ‘What privacy is for’ Harvard Law Review 126(7) pp 1904 - 1932

Criminal law - sentencing guidelines - Wisconsin Supreme Court requires warning before use of algorithmic risk assessments in sentencing - State v. Loomis. (2017, March). Harvard Law Review, 130(5), 1530+. Retrieved from https://link-gale-com.proxy.library.cmu.edu/apps/doc/A491136608/AONE?u=cmu_main&sid=AONE&xid=16a2ba98

Daniell, K, Morton A and Insua D 2016 ‘Policy analysis and policy analytics’ *Annals of Operations Research* 236:57-73

Danziger, S., Levav, J., and Avnaim-Pesso, L. (2011). *Extraneous factors in judicial decisions*. Proceedings of the National Academy of Sciences of the United States of America Vol. 108, pp 6889–6892

D.C. Law 22-111. Maternal Mortality Review Committee Establishment Act of 2018.
https://code.dccouncil.us/dc/council/laws/22-111.html

De Marchi G, Giulia Lucertini G, Tsoukiàs A 2016 ‘From evidence-based policy making to policy analytics’ *Annals of Operations Research* 236:15–38.

Dryzek, J and Stevenson H. "Global democracy and earth system governance." *Ecological economics* 70.11 (2011): 1865-1874.

Dryzek, John S. "Institutions for the anthropocene: governance in a changing earth system." *British Journal of Political Science* 46.4 (2016): 937-956

.

Dwork C, Hardty M, Pitassiz T, Reingold O, Zemel R November 30, 2011 ‘Fairness through Awareness’ Microsoft Research https://arxiv.org/pdf/1104.3913.pdf

Eaglin J M. 2017 ‘Constructing Recidivism Risk’. *Articles by Maurer Faculty. 2647.*
https://www.repository.law.indiana.edu/facpub/2647

Eubanks V 2018 *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. St Martin's Press. 2018.

Executive Office of the President 2016 *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*

https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

Freeman, K. (2016). "Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis." *North Carolina Journal of Law & Technology*, pp 88-89, 92

Fung, A and Wright, E 2003 'Thinking about Empowered Participatory Governance' in Fung, A. and Wright, E. (eds.) *Deepening Democracy*. Verso. New York

Gehman, Joel, et al. "Social license to operate: legitimacy by another name?" Canadian Public Administration, vol. 60, no. 2, 2017, p. 293+. Gale Academic OneFile

Green Paper for Vulnerable Children 2011. New Zealand Government https://img.scoop.co.nz/media/pdfs/1107/greenpaperforvulnerablechildren.pdf

Gunningham, Neil, Robert A. Kagan, and Dorothy Thornton. "Social license and environmental protection: why businesses go beyond compliance." *Law & Social Inquiry* 29.2 (2004): 307-341.

Helbing, D., Helbing and Caron, 2019. Towards Digital Enlightenment. Springer International Publishing.

House of Lords 2018 *AI in the UK: ready, willing and able?* https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf

Jackson, A. 2019. *Over 8,000 marijuana convictions in San Francisco dismissed with help from a computer algorithm* CNN February. https://edition.cnn.com/2019/02/25/us/san-francisco-marijuana-convictions-cleared-trnd/index.html

Jackson d and Marx G 2017 'Data mining pogram designed to predict child abuse proves unreliable, DCFS says' *Chicago Tribune* December 6 https://www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html

Janssen, Marijn, and George Kuk. 2016 "The challenges and limits of big data algorithms in technocratic governance." Government Information Quarterly Vol. 33 Issue 3. pp 371-377

Keddell, E. "The ethics of predictive risk modelling in the Aotearoa/New Zealand child welfare context: Child abuse prevention or neo-liberal tool?." *Critical Social Policy* 35.1 (2015):69-88.

Keddell, E., 2019. "Algorithmic Justice in Child Protection: Statistical Fairness, Social Justice and the Implications for Practice." *Social Sciences*, 8(10)

Kehl, D., Guo, P. and Kessler, S., 2017. “Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing.” Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School. 3. http://nrs.harvard.edu/urn-3:HUL.InstRepos:33746041

Kharpal, A 2020. *Google’s top policy chief calls for ‘common rules of the road’ globally for tech regulation* CNBC 20 February Google's policy chief calls for 'common rules' globally for tech regulation

Kleinberg J, Lakkaraju H, Leskovec J, Ludwig J & Mullainathan S 2017 ‘Human Decisions and Machine Predictions’ *The Quarterly Journal of Economics* Vol. 133(1), pp.237-293

Krantz, R 2003 ‘Cycles of Reform in Porto Alegre and Madison’ in Fung, A and Wright, E (eds.) *Deepening Democracy*. Verso. New York

Lenaghan J 1999 ‘Involving the public in rationing decisions. The experience of citizens juries’. *Health Policy* 49 45 - 61

Leonard, P 2018. Social licence and digital trust in data-driven applications and AI: a problem statement and possible solutions. Available at SSRN 3261228.

Maternal Mortality Review Committee 2019. Open DC. https://www.open-dc.gov/meeting/maternal-mortality-review-committee

McKee, M. "Citizens' juries: more questions than answers?." *Journal of Epidemiology & Community Health* 64.9 (2010): 750-751.

McKinsey 2011 *Big Data: The Next Frontier for Innovation, Competition and Productivity* https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation

McPhee, Gerard. "Centrelink's data-matching fiasco: A very predictable disaster." *Arena Magazine (Fitzroy, Vic)* 146 (2017): 20.

Masuga, H. 2018. ‘Toward a critical analysis framework of digital algorithms for policy makers.’ *Kennedy School Review*, 18, 104-108. Retrieved from https://search-proquest-com.proxy.library.cmu.edu/docview/2188098549?accountid=9902

Medhora R and Owen T 2020 ‘A Post-COVID-19 Digital Bretton Woods’ *Project Syndicate* April https://www.project-syndicate.org/onpoint/digital-bretton-woods-new-global-governance-model-by-rohinton-p-medhora-and-taylor-owen-2020-04

Mehr, H. 2017 *Artificial Intelligence for Citizen Services and Government*. Ash Center. Harvard University

Mintzberg, H. 1996 ‘Managing Government, Governing Management’. *HBR* May-June pp.75-83

Moore, M. 2013 *Recognising Public Value*. Harvard University Press.

Parsons, Richard, Justine Lacey, and Kieren Moffat. "Maintaining legitimacy of a contested practice: How the minerals industry understands its ‘social licence to operate’." *Resources Policy* 41 (2014): 83-90.

Price, David. "Choices without reasons: citizens' juries and policy evaluation." *Journal of Medical Ethics* 26.4 (2000): 272-276.

Putnam-Hornstein E and Needell, B. 2011 ‘Predictors of child protective service contact between birth and age five: An examination of California’s 2002 birth cohort’*. Children and Youth Services Review* 33 pp. 2400 – 2407

RNZ 2015 ‘Child abuse risk study 'ethically flawed' https://www.rnz.co.nz/news/national/280069/child-abuse-risk-study-'ethically-flawed'

Robbins, S., & Henschke, A. (2017). Designing for democracy: Bulk data and authoritarianism. Surveillance & Society, 15(3), 582-589. Retrieved from https://search-proquest-com.proxy.library.cmu.edu/docview/1929728949?accountid=9902

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), pp.206-215.

Rudin, C., Wang, C. and Coker, B., 2018. The age of secrecy and unfairness in recidivism prediction. *arXiv preprint arXiv:1811.00731*.

Sæbø, Ø., Rose, J., & Skiftenes Flak, L. (2008). ‘The shape of eParticipation: Characterizing an emerging research area.’ *Government Information Quarterly*, 25(3), 400–428. https://doi.org/10.1016/j.giq.2007.04.007

Scally, G 2018 *Scoping Inquiry into the Cervical Check Screening Programme* https://assets.gov.ie/9785/9134120f5b2c441c81eeed06808351c7.pdf

Stuff NZ National Health 2015 ‘Children 'not lab-rats' - Anne Tolley intervenes in child abuse experiment’ https://www.stuff.co.nz/national/health/70647353/children-not-lab-rats---anne-tolley-intervenes-in-child-abuse-experiment

Taylor, R 2020 ‘Pitt creates task force to eliminate bias in municipal algorithms’ *New Pittsburgh Courier*. January https://newpittsburghcourier.com/2020/01/27/pitt-creates-task-force-to-eliminate-bias-in-municipal-algorithms/

Tsoukias, A et. al. 2013 ‘Policy analytics: an agenda for research and practice’. *European Journal of Decisions Processes.* Vol. 1 pp 115 – 134

Vaithianathan, R 2012 ‘Can administrative data be used to identify children at risk of adverse outcomes?’ *Centre for Applied Research in Economics* (CARE) Department of Economics University of Auckland September

Vanian, J 2020 ‘Microsoft follows IBM and Amazon in barring police from using its facial-recognition technology’ Fortune June 12 https://fortune.com/2020/06/11/microsoft-ibm-amazon-facial-recognition-police/

White paper for Vulnerable Children 2012. Volume 1. New Zealand Government https://www.orangatamariki.govt.nz/assets/Uploads/white-paper-for-vulnerable-children-volume-1.pdf

Wilfond, B and Diekema. D 'Engaging Children in Genomics Research: Decoding the Meaning of Assent in Research. *Genetics in Medicine* 14.4 (2012): 437–443.

Williamson, B "Learning lessons from data controversies' *Code acts in education.* https://codeactsineducation.wordpress.com/2018/12/18/learning-lessons-from-data-controversies/

Zuckerberg, M 2020. 'Big tech needs more regulation' *Financial Times.* 17 February. https://www.ft.com/content/602ec7ec-4f18-11ea-95a0-43d18ec715f5