
Centre for Global Studies

Faculty Publications

The Limits of Policy Analytics: Early Examples and the Emerging Boundary of Possibilities

Justin Longo and Alan Rodney Dobell

2018

© 2018 the author(s), licensee Cogitatio (Lisbon, Portugal). This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0>), which permits any use, distribution, and reproduction of the work without further permission provided the original author(s) and source are credited.

This article was originally published at:

<http://dx.doi.org/10.17645/pag.v6i4.1561>

Citation for this paper:

Longo, J. & Dobell, A.R. (2018). The Limits of Policy Analytics: Early Examples and the Emerging Boundary of Possibilities. *Politics and Governance*, 6(4), 5-17.
<http://dx.doi.org/10.17645/pag.v6i4.1561>

Article

The Limits of Policy Analytics: Early Examples and the Emerging Boundary of Possibilities

Justin Longo^{1,*} and Rod Dobell²

¹ Johnson Shoyama Graduate School of Public Policy, University of Regina, Regina, S4S 0A2, Canada; E-Mail: justin.longo@uregina.ca

² Centre for Global Studies, University of Victoria, Victoria, V8P 5C2, Canada; E-Mail: rdobell@uvic.ca

* Corresponding author

Submitted: 1 May 2018 | Accepted: 19 September 2018 | Published: 21 November 2018

Abstract

Policy analytics has emerged as a modification of traditional policy analysis, where the discrete stages of the policy cycle are reformulated into a continuous, real-time system of big data collection, data analytics, and ubiquitous, connected technologies that provides the basis for more precise problem definition, policy experimentation for revealing detailed insights into system dynamics, and ongoing assessment of the impact of micro-scale policy interventions to nudge behaviour towards desired policy objectives. Theoretical and applied work in policy analytics research and practice is emerging that offers a persuasive case for the future possibilities of a real-time approach to policymaking and governance. However, policy problems often operate on long time cycles where the effect of policy interventions on behaviour and decisions can be observed only over long periods, and often only indirectly. This article surveys examples in the policy analytics literature, infers from those examples some characteristics of the policy problems and settings that lend themselves to a policy analytics approach, and suggests the boundaries of feasible policy analytics. Rather than imagine policy analytics as a universal replacement for the decades-old policy analysis approach, a sense of this boundary will allow us to more effectively consider the appropriate application of real-time policy analytics.

Keywords

adaptive management; agency; big data; data analytics; governance; nested institutions; nudging; policy analysis; policy analytics

Issue

This article is part of the issue “Big Data Applications in Governance and Policy”, edited by Sarah Giest (Leiden University, The Netherlands) and Reuben Ng (National University of Singapore, Singapore).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

If optimal control theory becomes fully operational in economics in the next few years...economists will have at their disposal a mathematical supertool that...actually tells you what policy to use...the best possible timing and dosage for each available policy remedy. (Business Week, 1973, p. 74)

1. Introduction

Policy analytics has emerged in recent years as a modification of the traditional policy analysis approach, where

the discrete stages of the policy cycle are being reformulated into a continuous, real-time system of big data collected from ubiquitous, connected technologies, assessed using advanced data analytics. Technological developments now provide policymaking with access to massive amounts of real-time data about policy problems and system conditions. When coupled with growing capacities in data analytics, policy analytics provides a basis for more precise problem definition, detailed insights into system dynamics, and ongoing assessment of the impact of micro-scale policy interventions to nudge

behaviour towards desired policy objectives (Daniell, Morton, & Insua, 2016; De Marchi, Lucertini, & Tsoukiàs, 2016; Höchtl, Parycek, & Schöllhammer, 2016; Kitchin, 2014; Lazer et al., 2009; Mergel, Rethemeyer, & Isett, 2016; Tsoukias, Montibeller, Lucertini, & Belton, 2013).

Policy analytics presents a mix of technology and expertise that could result in important advances in the science of policymaking (Giest, 2017). However, despite some early successes and enthusiasm for the possibilities of policy analytics, a number of questions and barriers to their use have emerged, principally issues related to privacy risks, data biases, and the need to clarify the relationship between the technocratic accuracy of policy analytics, and the challenges of decision-making in a diverse democracy (Noveck, 2018). Our focus here is on a specific concern that remains underexplored: to identify where the strengths of policy analytics live up to its billing, consider what the range of plausible applications is, and begin to assess the limits of policy analytics for addressing public policy problems. Our guiding research question asks what types of policy problems are amenable to ‘fast’ feedback control systems facilitated by big data and analytics, and which require a deeper, patient, ‘slower’ more deliberative approach to problem definition, analysis, decision-making, implementation, and evaluation (Kahneman, 2011). To pursue this question, we undertake a survey of the literature in policy analytics theory and practice, deriving from that the features of policy problems and their settings that characterize the range of policy issues to which policy analytics can reasonably be applied, leading towards a sketch of the boundaries of policy analytics. Rather than imagine policy analytics as a universal replacement for the decades old policy analysis approach, understanding this boundary will allow researchers and practitioners to more effectively consider the appropriate application of a real-time policy analytic approach. Our claim is that policy analytics complements and supports democratic deliberation and civic engagement; with agreement on operational objectives, policy analytics built on big data makes effective feedback control feasible.

We start by defining what we mean by policy analytics as distinct from policy analysis, sketch the emergence of the technological possibilities that have given rise to policy analytics and outline some concerns that have emerged. We next present a scan of recent policy analytic examples, leading to the identification of some characteristics of policy issues that are amenable to a policy analytics approach and—by extension—some types of policy issues that are not suitable to a continuous, real-time system of big data and data analytics, concluding with some guidance as to when policy analytics might be considered an appropriate approach. This boundary around the possibilities of policy analytics should supplement the broader need to consider the appropriate place for a policy analytic approach in the context of representative and deliberative democracy, social justice and equity considerations, social diversity, and citizen privacy

rights, concerns that should temper any unexamined enthusiasm for policy analytics.

2. The Emergence of Policy Analytics within the Policy Sciences

The modern policy analysis movement is based on an integrated, multidisciplinary approach to the study of public problems and the development of rational solutions based on careful analysis of evidence (Lerner & Lasswell, 1951). Decisions based on the best available evidence and rigorous analysis should be better positioned to address public problems than those based on anecdote, unsupported belief, or inaccurate data (Quade, 1975). From those origins, policy analysts have traditionally been tasked with precisely defining policy problems, collecting and analyzing data and evidence, supporting political decision-making with advice, guiding faithful implementation of those decisions, and objectively overseeing the evaluation of how effective those policy interventions were.

During the first quarter century of the policy analysis movement, quantitative techniques became staples of the theory and practice of policy analysis (Quade, 1980; Radin, 2000). Despite these significant advances and successes, debates over the perceived and proposed role of policy analysis have persisted in the profession’s later years (Dryzek, 1994; Stone, 1988). While technical, empirical, quantitative policy analysis became increasingly sophisticated during the 1970s, and since, high-profile failures and the perceived inability to solve complex public problems exposed the limits of positivist policy analysis (May, 1992). Critics of positivism argued that the attempt to model social interactions using mathematical models was misguided (Amy, 1984), that policy analysis was much more than data analysis (Meltsner, 1976; Wildavsky, 1979), and that positivism was fundamentally incapable of dealing with complex problems in a democracy (Fischer, 1995). A “malaise...of the policy sciences” crept into the discipline as its positivist, neo-classical economics orientation seemed incapable of understanding human behaviour, accommodating the democratic expectations of citizens, or remedying the increasing complexity of policy problems (Deleon, 1994, p. 82). The positivist policy analysis hegemony was also undermined by limitations in data availability and the tools of analysis (Morgan, Henrion, & Small, 1992). Analysts inclined towards quantitative methods longed for even more robust data, greater computational power, and the development of more technically sophisticated policy analysis throughout government and wider policy circles (Morçöl, 2001). Some of those goals appear to have been attained in the digital era, with the growth of big data arising from the ubiquitous deployment of networked computing devices throughout society and increased data analytic capacity to manage the resulting flood of data.

Definitions of ‘big data’ abound (Dutcher, 2014; Fredriksson, Mubarak, Tuohimaa, & Zhan, 2017; Ward &

Barker, 2013), with most focusing on its characteristics—especially the large *volume* of data, its continuous flow at high *velocity*, and the *variety* of data available—and others pointing to the complexity of combined data sets and their value in revealing previously undetectable patterns. What emerges, however, from the policy analytics literature is a frequent conflation of ‘big data’ with ‘large’ data collections such as a census. While this reflects the current state of the art, our concept of big data draws especially on the velocity and variety (and, consequently, the large volume) of data as the foundation for a policy analytic approach that centres on a real-time understanding and interaction with the policy environment.

With the emergence and expansion of the Internet and the range of digital technologies that have been deployed in recent years, analysts now have access to a wide range of policy-relevant big data. These technologies and their users generate a variety of signals through devices like mobile smartphones, Internet of Things (IoT) devices, personal wearables, electronic transaction cards, *in situ* sensors, web search and web traffic, and social media. Massive amounts of data are now generated continuously through the daily activities of individuals, from their interactions with web services and social media platforms, purchasing behaviour and transportation choices revealed through electronic transaction cards, movement and interaction captured through mobile smartphones and wearables, behavioural choices measured through IoT consumer products, a range of measurements captured by sensors, satellite remote sensing, counters, and smart meters, and the interactions of people and devices with control technology. The accumulation of these signals, and associated metadata such as geolocation information and time stamps, results in a previously unimaginable amount of data, precisely measured from multiple perspectives, and captured in real time. Advances in data storage technologies now make it possible to preserve increasing amounts of data, and faster data transfer rates allow for cloud computing at low cost. We can now capture, store, and process data—in volumes previously unimaginable, from ubiquitous sources, with continuous flow, observed through multiple channels—and have increased capacity to manage, analyze, and understand these new data sources (Lazer et al., 2009). Not only has the volume of data and our ability to analyze it changed. The same technologies that allow for real-time data capture from the field provide a mechanism to communicate policy signals outward to actors, agents, and those devices, serving again to gather further data that measure reaction to those signals. With the stages thus joined up, policy formulation can be connected with implementation and evaluation processes in a continuous and real-time cycle of ideation, experimentation, evaluation, and reformulation (Pirog, 2014). New digital tools, platforms, and the data they generate allow for a seamless linking of the discrete stages of the policy cycle into a continuous, real-time, feedback cycle of problem identification, tool mod-

ification, system monitoring, and evaluation. This technology revolution offers the potential to revive and extend the positivist tradition in policy analysis and offer improved support for policymaking through an approach we call ‘policy analytics’.

To be certain, there are competing conceptualizations of what policy analytics implies (Daniell et al., 2016; De Marchi et al., 2016; Tsoukias et al., 2013). While referred to *inter alia* as ‘big data’ applied to public policy and administration (Einav & Levin, 2014a; Giest, 2017; Höchtl et al., 2016; Kim, Trimi, & Chung, 2014; Kitchin, 2014; Mergel et al., 2016), ‘computational social sciences’ (Lazer et al., 2009), and ‘policy informatics’ (Johnston, 2015), the term policy analytics is used here to emphasize the combination of new sources and forms of policy-relevant big data with the use of new analytic techniques and capacity to affect policymaking throughout the entire policy cycle. Some definitions stretch the definition of ‘big data’ to include traditional—albeit very large—government ‘large data’ collections such as censuses, taxation data, social security records, health information, and survey data (Daniell et al., 2016). Some perspectives emphasise this supplementing of large data with big data, where datasets are linked with the aim of identifying previously undiscovered patterns and correlations at the problem identification and analysis stages (Höchtl et al., 2016; Janssen & Kuk, 2016a). Others focus on high volume real-time big data, combined with highly structured administrative large data, for deriving insights for operations and public service delivery (Joseph & Johnson, 2013; Mergel et al., 2016).

The harvesting of big data, coupled with advances in technology and scientific developments for managing that data, emerged first in the private sector under the heading ‘business analytics’, with analytics serving as an umbrella term for statistical methods and approaches including statistics, data mining, machine learning, business intelligence, knowledge management, decision support systems, operations research, and decision analysis. Key to the development of business intelligence was that this intelligence was useful if it led to action that was immediate and the impact of that action measurable (Longo, 2018; McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012). When eventually applied to public policy problems, this led to the concept of ‘policy analytics’ denoting the development and application of data analytic skills, methods, and technologies, supporting stakeholders with meaningful and informative analysis at any stage of the policy cycle (De Marchi et al., 2016; Tsoukias et al., 2013). Pirog (2014) envisions the extension of previously developed quantitative methods through the linking of government administrative records, data from natural science fields such as biology and neuropsychology, and geospatial data ushering in a dramatic advance in policy research. Giest (2017) gives examples from different policy domains—health, education, climate change, and crisis management—and identifies a mix of data, technology, and expertise that could result in important ad-

vances in the science of policymaking. Thus, based on the literature that has emerged to date from both business analytics and policy analytics, we define policy analytics as the use of new sources and forms of policy-relevant big data combined with advanced analytics techniques and capacity, taking advantage of ubiquitous communication methods to reduce the time delay amongst stages of the policy cycle, aimed at better addressing public problems.

In adopting the tools of policy analytics, governments are mirroring the actions of private sector firms that use big data to better understand people's behaviour. Examples include encouraging users to return to a webpage, click on an ad, buy a product and a subsequent product, purchase a service, or watch a movie because they watched a similar one (McAfee et al., 2012). Data analytics can also be used to judge who is a worthy credit risk, who would be a good person to hire, and who would make an ideal mate (Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016; Tufekci, 2014). Despite these early successes and enthusiasm for the possibilities of policy analytics, a number of questions and barriers to their use have emerged that should temper any unexamined enthusiasm (Noveck, 2018). Among these are concerns over privacy and security of citizens' data (Kim et al., 2014), proper and efficient permissioning to facilitate use by public servants (Welch, Hinnant, & Moon, 2005), weak data skills among public servants and a reliance on external consultants and contract data analysts (Dunleavy, Margetts, Bastow, & Tinkler, 2006), faulty analysis where strong correlations are valued over preliminary causal explanations (Harford, 2014), questions about big data representativeness as new digital divides emerge that undermine the possible democratizing effects of policy analytics (Longo, Kuras, Smith, Hondula, & Johnston, 2017), establishing the provenance of big data so that stakeholders and decision makers can understand where the evidence came from (Javed, McClatchey, Khan, & Shamdasani, 2018), opacity in policymaking by algorithm (Kitchin, 2017; Mittelstadt et al., 2016; Pasquale, 2015), bias in algorithms and machine learning (Koene, 2017), an over-reliance on data for decision-making in situations where values are important (Majone, 1989; Shulock, 1999), and its inverse, ignoring data in decision-making (Harsin, 2015; Tingling & Brydon, 2010).

Policy analytics represents a persuasive combination of advanced digital technology and modern behavioural science. But it has emerged alongside volatile and untrustworthy information and communications technologies reshaping shifting perceptions and redirecting changing beliefs, driving the evolving preferences that must be reflected in contested metrics for signalling social welfare and community wellbeing. In assessing this challenge, it is necessary to consider what kinds of public reasons can legitimately support the authoritative exercise of delegated public power in a political setting marked by a lack of consensus within a divided society.

As the potential dangers of the big data industry begin to be revealed and slowly understood (Persily, 2017), the question that must be asked of government is whether the benefits of policy analytics outweigh the potential downsides (Boyd & Crawford, 2012). This challenge is, of course, just one facet of the broader social question of what it means to retain meaningful human control of technocratic instruments, including autonomous and intelligent systems, in a world where the exercise of human agency is increasingly distanced from consequences and individual responsibility.

3. Policy Analytics in Practice

Policy analytics can take a range of approaches. Perhaps the simplest, first line of analysis lies in social media monitoring or 'social listening' to analyze and respond to citizen's preferences, experiences, articulated values, and behaviours (Charalabidis, Loukis, Androutopoulou, Karkaletsis, & Triantafillou, 2014; Grubmüller, Götsch, & Krieger, 2013; Prpić, Taeihagh, & Melton, 2015). Social listening involves searching and monitoring social media for words, phrases, hashtags, or mentions of government accounts or persons. This approach is becoming increasingly popular with governments seeking to gauge public perception and better appreciate why citizens have the attitudes they do and how these attitudes change over time (Longo, 2017; Paris & Wan, 2011). Further analysis can centre on the assessment of sentiment and meaning, clustering opinion to reveal network properties and make sense of public opinion (Till, Longo, Dobell, & Driessen, 2014).

Venturing deeper, predictive analytics can serve as an input into framing a policy problem before it is apprehended as such, indicating where a need is being unmet, or where an emerging problem might be countered early. As a big data analytics form of forecasting (Sims, 1986) now referred to as nowcasting (Choi & Varian, 2012), predictive analytics is based on the argument that analysis of past performance can reveal a probable outcome that can be expected from continuing to pursue the same approach (i.e., doing nothing). Some recent initiatives show the possibilities for success, for example in reducing administrative failures (Behavioural Insights Team, 2012) and understanding social dynamics (Bond et al., 2012). The combination of digital signals and new analytic techniques can help in understanding and predicting behavior in contexts such as crime (Chan & Bennett Moses, 2015), energy use (Zhou & Yang, 2016), migration (e.g., the use of email, social media, web search, and geolocation have been used to infer migration flows; see Gerland, 2015; Raymer, Wiśniowski, Forster, Smith, & Bijak, 2013; Verhulst & Young, 2018), urban planning (Kitchin, 2014), and public health (Khoury & Ioannidis, 2014; Murdoch & Detsky, 2013).

Policy experimentation builds on the idea of policy incrementalism (Lindblom, 1959), with a long history of examples of trials, experiments, and pilots of varying

scale and precision, and a renewed enthusiasm in jurisdictions from the United Kingdom (Breckon, 2015) to Canada (Monafu, Chan, & Turnbull, 2018). Real-time experimental policy analytics takes advantage of new big data sources, coupled with data analytics techniques, bringing together all the discrete stages of the policy cycle into one continuous process. While a policy problem is being observed, interventions would also be underway using the same devices used to collect the data, with their impact on the problem becoming part of the evidence base for further modifying the policy variables. These further modifications would also be observed for their impact, as the system response to the policy intervention moved closer to the policy target or equilibrium (Esperanza & Dirk, 2014). An intriguing application of policy analytics from transportation management can be seen in the evolution from high-occupancy vehicle (HOV) lanes to high-occupancy smart toll (HOST) lanes (Longo & McNutt, 2018).

Shi, Ai and Cao (2017) argue that some policy analytic methods are better suited to particular stages of the policy cycle than others, and provide several examples to support their claim. Cognitive mapping, text mining, and understanding public attitudes through geo-specific Google search-query data (Lee, Kim, & Lee, 2016) are applicable to the agenda-setting phase, participatory planning in the decision phase, and remote sensing, smart metering, or participatory GIS to monitoring and evaluation phases. Decision support systems to collect, manage, and analyze data (e.g., a space-air-ground big data traffic system that includes people, vehicles, and road conditions using data from satellite sensing, aerial photography, aerial drone sensors, cameras, transponders, and smartphones) can support overall transportation policy implementation, law enforcement, and emergency response (Xiong et al., 2016). A groundwater web portal that combines legacy data, community-sourced groundwater information, and government open data provides real-time information to the public, and tools for data querying and visualization to support decision-making and community engagement (Dahlhaus et al., 2016). A big data archive covering more than 43 million soldiers, veterans, and their family members provides a foundation for the examination of the causes and consequences of PTSD (Vie, Griffith, Scheier, Lester, & Seligman, 2013).

In some cases, policy evaluation can be undertaken using policy analytics. Lu, Chen, Ho and Wang (2016) analyze 2 million construction waste disposal records to assess the disparity between public and private operator performance, with contractors operating in public projects performing better than those in private projects. In transportation management, cases from the Netherlands and Sweden show that automated smart-card and vehicle positioning data provide for better understanding of passenger needs and behaviours, system performance, and real-time conditions in order to support planning and operational processes (Van Oort & Cats, 2015).

Participatory policy analytics can take the form of sentiment analysis, mined from Internet content including social media, used to gauge how the public values alternative outcomes. Beyond simplistic exercises such as counting 'likes' and 'mentions', the example of mining Yelp restaurant reviews as a supplement (and potential replacement) for public health inspections (Kang, Kuznetsova, Luca, & Choi, 2013) shows that mining of large volumes of text contributions from citizens concerning government policies can extract opinions and knowledge useful for policy purposes (Maragoudakis, Loukis, & Charalabidis, 2011).

Poel, Meyer and Schroeder (2018) present the results of a recent project that scanned for big data policymaking examples, noting the heightened interest in big data for policymaking in recent years, though acknowledging that there are still few good examples available. They analyze 58 data-driven cases, with a focus on national and international policy initiatives, and highlight persistent challenges: data representativeness, validity of results, gaps in citizen engagement, and weak data analysis skills. While most examples do not tread on personally identifiable data, privacy protection remains a concern due to re-identification/de-anonymization risks (de Montjoye, Hidalgo, Verleysen, & Blondel, 2013; Narayanan & Shmatikov, 2008). More generally, using big data for policy purposes revives concerns about technocracy, technoscience, policy-based evidence making, and the influence of lobby groups. The most prominent area Poel et al. (2018) identify centres on government transparency initiatives supported by the publication of open data on procurement, having the objective of revealing government corruption. A smaller number of initiatives focus on operational policy areas such as budgeting, economic and financial affairs, and transportation. Remaining initiatives cover policy areas such as health, education, research, justice, and social affairs. Almost half of the initiatives scanned focus primarily on the early stages of the policy cycle (e.g., sentiment analysis via Twitter to support agenda setting and problem analysis), with others supporting policy design, implementation, and monitoring. Observing traffic patterns via sensors and mobile phone data, and using administrative data to monitor transportation and environmental policies, were also highlighted. However, as most of the projects scanned in Poel et al. (2018) use data formats such as spreadsheets, and data analysis is limited to descriptive statistics or occasional visualizations with few examples of techniques such as machine learning or algorithmic response, the boundary in this survey between 'large data' and 'big data' appears fluid.

Schintler and Kulkarni (2014) review the range of arguments for and against the use of big data in policy analysis, and offer examples to illustrate some of the positive features. The 'Billion Prices Project' uses web-sourced price information from retailers across multiple countries and sectors to generate daily estimates of inflation, providing a real-time price index as opposed to

the periodic figures produced by national statistical agencies (Cavallo & Rigobon, 2016). The 'Global Forest Watch' project processes hundreds of millions of satellite images as well as data from people on the ground to generate real-time estimates of tree loss that are more precise than those produced from other approaches (Hartmann et al., 2018). The near real-time data are available freely online, and have been used to measure global deforestation rates, detect illegal clearing activity and burning, and monitor corporate sustainable forestry commitments.

Daniell et al. (2016) point towards examples of policy analytics for formulation or delivery in the areas of health resource allocation (Aringhieri, Carello, & Morale, 2016), sentiment analysis and opinion mining (Alfaro, Cano-Montero, Gómez, Moguerza, & Ortega, 2016), using behavioral information to encourage energy efficiency, precision government services (Hondula, Kuras, Longo, & Johnston, 2017), identifying social service and public information 'deserts' (Entwistle, 2003), and promoting smart cities (Kumar, Nguyen, & Teo, 2016). Additional examples are being tested, and stand as potential opportunities for applied policy analytics, from using smart electricity meters to incentivize conservation behaviour and reduce peak-load demand (Blumsack & Fernandez, 2012; Newsham & Bowker, 2010), to possibilities such as creating on-demand local public transportation services (Murphy, 2016). The Joint Statistical Research Program of the US Internal Revenue Service enables studies that use long panels of tax returns to observe individuals over time with a view to revealing potential policy initiatives (Jarmin & O'Hara, 2016).

The principles of nudge theory are being applied in dynamic ways that take advantage of the powerful devices ubiquitously moving around us to measure the environment, along with individual behavior and health conditions, to intervene by sending information to the individual via devices such as their smartphone in order to change a behavior (Katapally et al., 2018). Smart devices can be deployed to monitor behaviour in teams to improve performance (Pentland, 2012), or monitoring student engagement to improve learning outcomes (Crosby & Ikehara, 2015).

The recent advances in Artificial Intelligence (AI) that we are currently experiencing—e.g., autonomous vehicles, facial recognition—have accelerated due to the combined developments of big data and analytics, especially machine learning. However, the origins of AI, and concerns over its adoption in public policy and administration, are much deeper. The early promise of AI in public sector practice centred on providing decision support for public managers (e.g., Barth & Arnold, 1999; Hadden, 1986; Hurley & Wallace, 1986; Jahoda, 1986; Masuch & LaPotin, 1989) but failed to materialize in any meaningful way. While the early promises of AI went unfulfilled, there have been dramatic advances in AI in recent years (Russell & Norvig, 2009) that could have important consequence for public management and governance. A key contributing factor to increasing maturity

of AI technologies and the viability of AI application to public policy and administration is the increased availability of data that can be used to further machine learning. As algorithms become more widely used, increasingly autonomous, and invisible to those affected by their decisions, their status as impartial public servants becomes more difficult to monitor for bias or discrimination (Janssen & Kuk, 2016b). Today, AI systems are being used to detect irregularities, with aims such as reducing fraud and errors in service processing (Maciejewski, 2017). An even more speculative example (Death, 2015) addresses challenges of watershed governance, envisaging the application of AI to the continuous monitoring of complex streamflow dynamics and water chemistry and quality as part of decision support systems for communities concerned with environmental flows as well as crucial water supply. The possible extension to Artificial Intelligence that could offer, autonomously, better decisions than the community might make in resolving the conflicts around the vital tradeoffs among the many interests, human and ecological—as well, perhaps, as the rights of the river itself—is a topic of ongoing debate. Relatedly, the question of meaningful human involvement in decisions related to problems of human security has been addressed in a recent report on the role of AI in nuclear war (Geist & Lohn, 2018).

4. Discussion

Given the scan of examples of policy analytics in practice, where does this revision to the policy analysis model fit in the modern governance toolkit, and what do the examples of successful policy analytics applications tell us about the possibilities for its future, and the limitations it will likely face?

We must be careful not to overstate what policy analytics can tell us. Take, for example, the rhetoric around predictive analytics (Gandomi & Haider, 2015), which can serve as an input into framing a policy problem before it is apprehended as such. In 'predictive policing', where potential crimes, offenders, and victims are identified *a priori*, police resources can be directed proactively (Brayne, 2017; Perry, McInnis, Price, Smith, & Hollywood, 2013). The inherent complexities of social, economic, and behavioural phenomena, however, make policy prediction essentially impossible (Sawyer, 2005). While modeling for purposes of forecasting (Sims, 1986) and related approaches such as backcasting (Robinson, 1982) can serve as useful tools in policy analysis, and these techniques have improved with the increase in available data and growth in analytical capacity (Einav & Levin, 2014b; Wang, Kung, & Byrd, 2018), there are obvious limits on our ability to predict the future. Predictive models are necessarily abstractions from reality, and cannot feasibly include all individual and system factors. More likely are qualitative statements (including probability statements as to likelihood) about the direction of predicted change, including indications about possible

unexpected outcomes. These are useful for policy analysis, especially in highly uncertain environments where unlikely events may still yield catastrophic outcomes.

It should be obvious that the proposed policy analytic approach will not solve all policy challenges. Despite the power of modern digital technology, a number of limitations and caveats remain. While more, and more accurate, evidence can improve our understanding and form the basis for better policy, we should not conflate the volume of big data with its representativeness. Despite the mesh of sensors that act as the collection net for policy-relevant data, there is the risk that those without the right devices or engaged in the targeted behavior may be rendered “digitally invisible” in the movement towards rapid policy design (Longo et al., 2017, p. 76). There are also a number of technical limits to assembling robust big data sets including challenges in data acquisition (especially where much of the really valuable data is closely guarded by private companies; (Golder & Macy, 2014; Verhulst & Young, 2018), data interoperability problems (Miller & Tucker, 2014), and legitimate privacy protections that place prohibitions on the sharing of data outside of programs or departments, or even on combining datasets behind protective firewalls. Even if data coverage is comprehensive, big data hubris can produce policy errors (Lazer, Kennedy, King, & Vespignani, 2014). Traditional social science designs research instruments to collect data in order to test a hypothesis, whereas big data analytics seeks to identify relationships (Wigan & Clarke, 2013). And the risk of apophenia—the seeing of patterns in random data—can lead policymakers to identify correlations that are easily mistaken for causal relationships (Boyd & Crawford, 2012).

Shi et al. (2017, p. 552) note that “only a few government decisions have already benefited from the systematic use of masses of data and evidence, and of cutting-edge modelling”, with the norm being to rely on traditional forms of policy analysis. Several challenges are noted, centring on the democratic underpinnings of policy analysis. Since public sector problems typically involve making decisions on behalf of society at large, involving the allocation of public resources (Lerner & Lasswell, 1951), policy analytics must balance the need for robust analysis with the need to satisfy legitimacy expectations, transparency requirements, and opportunities for citizen participation.

Thus the policy analytics model—of the rapid prototype based on a digitally enabled system of communication, feedback, analytics and tool modification—does not apply across a wide range of policy problems or domains. Many policy areas are not amenable to minor policy tool modifications that can be communicated digitally. Few policy systems form such a tight linkage between a minor modification of a policy signal and an immediately detectable response from the system under observation, instead operating across long timescales between policy intervention and system response. Policy analytics is well suited to the digital realm of approaches such as A/B

testing of government citizen service websites (Longo, 2018), whereas many policy decisions entail actions that have significant consequences diffused over many sectors. More often than not, the policy environment will be complex beyond the capabilities of even the most advanced analytics. The possibility of policy experimentation will apply in a limited set of circumstances, especially where legitimate ethical concerns could be raised.

Consider the 4-quadrant diagram in Figure 1, with the horizontal axis running from micro or local scale on the left through regional or meso-scale to global scale on the right, and the vertical axis from certainty as to system structure and environment at the bottom to profound uncertainty at the top. In the top right quadrant (high uncertainty, global scale) one has ‘wicked problems’, ‘messes’, concerns of post-normal science facing all the challenges of affective conflict and democratic dissent. Examples might be climate change, global hydrological cycle, poverty and inequality. But even in these challenging settings one can look to rapidly increasing computational capacity to develop decision support systems. To the extent that agreement can be achieved on appropriate policy objectives and instruments, there can be realistic ambitions for real-time policy analytic systems.

In the top left quadrant, more inclusive community engagement and deliberation, building on increasingly sophisticated decision-support systems, is feasible, but again expectations of integrated data analytic/policy analytic systems running on a real-time basis must rest on hopes for inclusive and collaborative policy formation processes building agreement on legitimate and acceptable policy objectives and norms of implementation. The lower right quadrant might be thought largely empty for the moment: there appear to be few global scale challenges for which one can have reasonable certainty around system structure and environment, except perhaps international agreements on classification systems or the like. But even here, as international agreements grow in number and specificity, policy analytic methods for monitoring and certifying compliance are increasingly significant.

Nevertheless, it seems that it is in the lower left quadrant, with reasonable certainty around the nature and context of micro or local scale problems that big data, data analytics and policy analytics can best support ongoing experimentation, continuous learning, policy formation, and adaptive management with effective implementation, monitoring and enforcement. Focusing on this quadrant, how might its boundaries evolve and expand? Evidently the operational problems faced in managing the direct provision of services at local level are more amenable to such experimentally-based adaptive control and self-regulation than for the problems that have to be addressed through cooperative federalism or similar institutional arrangements for negotiation among authoritative political units at different scales. Although the professional effort to differentiate the ‘policy design’ product from the more traditional language calls

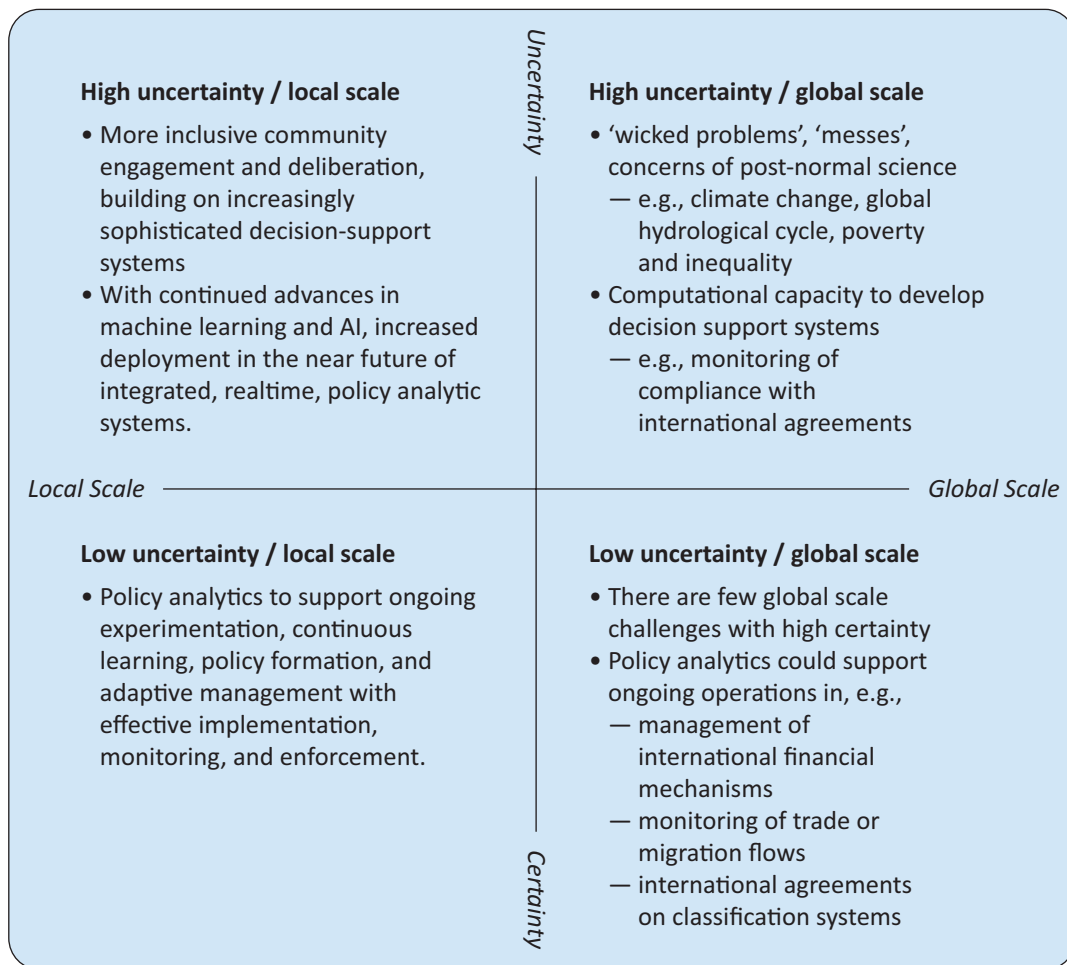


Figure 1. The applicability of policy analytics across scale and complexity.

attention more to the implementation end of the cycle than to the formulation portion, the bigger challenge for the rapid adaptation of design in response to user experience lies in the varied and slow instruments for implementing change in the operations of representative government, legitimately and with ongoing accountability. The fuzzy boundaries that separate a summative evaluation cycle for Cabinets or executive authorities from a formative evaluation cycle for management exercising delegated authority in decisions at small (how small?) scale might suggest limits to policy analytics—but they also suggest the potential of machine learning and autonomous and intelligent systems in pushing those boundaries far outward. The science fiction aspects of Joe AI, analyst, or Jane AI, authoritative decision-maker—and the challenges of teaching her/it in new schools of public policy—may be with us much sooner than expected, with consequent rapid advance in the spread of policy analytics as integrated system.

5. Conclusion

This article began with a quotation from a leading business magazine in 1973 that enthused about the possibilities of a policy supertool that then appeared immi-

nent. That quotation was cited in a commentary from the Honourable C. M. Drury (then President of the Treasury Board of Canada—the agency charged with the development of tools for policy analysis and decision support in the Government of Canada at the time) in the inaugural issue of the journal *Canadian Public Policy*. In reaction to the fantastic possibilities envisaged, the Minister suggested that “While we may all have our occasional doubts about the advice offered by our traditional public servants, I am certainly not yet ready to trade them in on the strength of this promise!” (Drury, 1975, p. 91).

Almost a half-century later, does policy analytics represent the delayed realisation of that promised policy supertool, or yet another misplaced enthusiasm? Daniell et al. (2016, p. 11) conclude their special issue of policy analytics in practice with the consideration “that analytics have been somehow oversold”, that political decision making can be overcome by masses of data, and deep analytics, producing automated solutions to any public problem. While evidence is important, decision making still requires judgment. New initiatives can be informed by past experience, but still require careful experimentation to avoid large implementation failures.

The emerging examples may be persuasive in their particular domains. But many of the problems con-

fronted by policy analysts are indeed wicked problems involving differing time scales in complex systems where the effects of policy interventions on decisions and behaviour are unclear, uncertain, and of unknown duration. Much more crucially, agreement on the objectives or purposes of policy is usually lacking, and interests around the nature or instruments of policy intervention are conflicted. Not all policy environments are compatible with the policy analytics model. Much work remains to be done before we find the proper place for this promising development in an increasingly post-positivist, post-fact, post-truth world.

Acknowledgements

The comments of two anonymous reviewers, and the managing and academic editors for this thematic issue, are greatly appreciated.

Conflict of Interests

The authors declare no conflict of interests.

References

- Alfaro, C., Cano-Montero, J., Gómez, J., Moguerza, J. M., & Ortega, F. (2016). A multi-stage method for content classification and opinion mining on weblog comments. *Annals of Operations Research*, 236(1), 197–213.
- Amy, D. J. (1984). Towards a post-positivist policy analysis. *Policy Studies Journal: The Journal of the Policy Studies Organization*, 13(1), 207–211.
- Aringhieri, R., Carello, G., & Morale, D. (2016). Supporting decision making to improve the performance of an Italian emergency medical service. *Annals of Operations Research*, 236(1), 131–148.
- Barth, T. J., & Arnold, E. (1999). Artificial intelligence and administrative discretion: Implications for public administration. *American Review of Public Administration*, 29(4), 332–351.
- Behavioural Insights Team. (2012). *Applying behavioural insights to reduce fraud, error and debt*. London: Cabinet Office. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60539/BIT_FraudErrorDebt_accessible.pdf
- Blumsack, S., & Fernandez, A. (2012). Ready or not, here comes the smart grid! *Energy*, 37(1), 61–68.
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415), 295–298.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication and Society*, 15(5), 662–679.
- Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, 82(5), 977–1008.
- Breckon, J. (2015). *Better public services through experimental government*. London: Alliance for Useful Evidence. Retrieved from <http://www.capire.org/capireinforma/scaffale/Final.pdf>
- Business Week. (1973, May 19). Optimal control: A mathematical supertool. *Business Week*.
- Cavallo, A., & Rigobon, R. (2016). The billion prices project: Using online prices for measurement and research. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 30(2), 151–178.
- Chan, J., & Bennett Moses, L. (2015). Is big data challenging criminology? *Theoretical Criminology*, 20(1), 21–39.
- Charalabidis, Y., Loukis, E. N., Androutsopoulou, A., Karkaletsis, V., & Triantafillou, A. (2014). Passive crowdsourcing in government using social media. *Transforming Government: People, Process and Policy*, 8(2), 283–308.
- Choi, H., & Varian, H. (2012). Predicting the present with Google trends. *The Economic Record*, 88, 2–9.
- Crosby, M. E., & Ikehara, C. S. (2015). Feedback from physiological sensors in the classroom. In T. Hammond, S. Valentine, A. Adler, & M. Payton (Eds.), *The impact of pen and touch technology on education* (pp. 381–387). Cham: Springer International Publishing.
- Dahlhaus, P., Murphy, A., MacLeod, A., Thompson, H., McKenna, K., & Ollerenshaw, A. (2016). Making the invisible visible: The impact of federating groundwater data in Victoria, Australia. *Journal of Hydroinformatics*, 18(2), 238–255.
- Daniell, K. A., Morton, A., & Insua, D. R. (2016). Policy analysis and policy analytics. *Annals of Operations Research*, 236(1), 1–13.
- Death, R. G. (2015). An environmental crisis: Science has failed; let us send in the machines. *Wiley Interdisciplinary Reviews: Water*, 2(6), 595–600.
- Deleon, P. (1994). Reinventing the policy sciences: Three steps back to the future. *Policy Sciences*, 27(1), 77–95.
- De Marchi, G., Lucertini, G., & Tsoukiàs, A. (2016). From evidence-based policy making to policy analytics. *Annals of Operations Research*, 236(1), 15–38.
- de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3, 1376.
- Drury, C. M. (1975). Quantitative analysis and public policy making. *Canadian Public Policy. Analyse de Politiques*, 1(1), 89–96.
- Dryzek, J. S. (1994). *Discursive democracy: Politics, policy, and political science*. Cambridge: Cambridge University Press.
- Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New public management is dead—Long live digital-era governance. *Journal of Public Administration Research and Theory*, 16(3), 467–494.
- Dutcher, J. (2014, September 3). What is big data?

- datascience@berkeley*. Retrieved from <https://datascience.berkeley.edu/what-is-big-data>
- Einav, L., & Levin, J. (2014a). Economics in the age of big data. *Science*, *346*(6210), 1243089.
- Einav, L., & Levin, J. (2014b). The data revolution and economic analysis. *Innovation Policy and the Economy*, *14*, 1–24.
- Entwistle, V. (2003). Editorial. *Health Expectations*, *6*(2), 93–96.
- Esperanza, L., & Dirk, P. (Eds.). (2014). *Making innovation policy work: Learning from experimentation*. Paris: OECD Publishing.
- Fischer, F. (1995). *Evaluating public policy*. Chicago: Nelson-Hall.
- Fredriksson, C., Mubarak, F., Tuohimaa, M., & Zhan, M. (2017). Big data in the public sector: A systematic literature review. *Scandinavian Journal of Public Administration*, *21*(3), 39–62.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144.
- Geist, E., & Lohn, A. J. (2018). How might artificial intelligence affect the risk of nuclear war? *RAND Corporation*. Retrieved from <https://www.rand.org/pubs/perspectives/PE296.html>
- Giest, S. (2017). Big data for policymaking: Fad or fast-track? *Policy Sciences*, *50*(3), 367–382.
- Gerland, P. (2015). *Migration, mobility, and big data: An overview*. New York, NY: United Nations.
- Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology*, *40*, 129–152.
- Grubmüller, V., Götsch, K., & Krieger, B. (2013). Social media analytics for future oriented policy making. *European Journal of Futures Research*, *1*(1), 20.
- Hadden, S. G. (1986). Intelligent advisory systems for managing and disseminating information. *Public Administration Review*, *46*, 572–578.
- Harford, T. (2014). Big data: A big mistake? *Significance. Statistics Making Sense*, *11*(5), 14–19.
- Harsin, J. (2015). Regimes of post-truth, post politics, and attention economies. *Communication, Culture and Critique*, *8*(2), 327–333.
- Hartmann, H., Schuldt, B., Sanders, T. G. M., Macinnis-Ng, C., Boehmer, H. J.S., Allen, C. D., . . . Anderegg, W. R. L. (2018). Monitoring global tree mortality patterns and trends. Report from the VW symposium ‘Crossing scales and disciplines to identify global trends of tree mortality as indicators of forest health’. *New Phytologist*, *217*, 984–987.
- Höchtel, J., Parycek, P., & Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, *26*(1/2), 147–169.
- Hondula, D. M., Kuras, E. R., Longo, J., & Johnston, E. W. (2017). Toward precision governance: Infusing data into public management of environmental hazards. *Public Management Review*, 1–20.
- Hurley, M. W., & Wallace, W. A. (1986). Expert systems as decision aids for public managers: An assessment of the technology and prototyping as a design Strategy. *Public Administration Review*, *46*, 563–571.
- Jahoda, M. (1986). Artificial intelligence: An outsider’s perspective. *Science & Public Policy*, *13*(6), 333–340.
- Janssen, M., & Kuk, G. (2016a). Big and open linked data (BOLD) in research, policy, and practice. *Journal of Organizational Computing and Electronic Commerce*, *26*(1/2), 3–13.
- Janssen, M., & Kuk, G. (2016b). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, *33*(3), 371–377.
- Jarmin, R. S., & O’Hara, A. B. (2016). Big data and the transformation of public policy analysis. *Journal of Policy Analysis and Management*, *35*(4), 715–721.
- Javed, B., McClatchey, R., Khan, Z., & Shamdasani, J. (2018). A Provenance framework for policy analytics in smart cities. *arXiv [cs.CY]*. Retrieved from <http://arxiv.org/abs/1804.07141>
- Johnston, E. W. (2015). *Governance in the information era: Theory and practice of policy informatics*. New York, NY: Routledge.
- Joseph, R. C., & Johnson, N. A. (2013). Big data and transformational government. *IT Professional*, *15*(6), 43–48.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Kang, J. S., Kuznetsova, P., Luca, M., & Choi, Y. (2013). Where not to eat? Improving public policy by predicting hygiene inspections using online reviews. In *proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1443–1448). Stroudsburg, PA: aclweb.org.
- Katapally, T. R., Bhawra, J., Leatherdale, S. T., Ferguson, L., Longo, J., Rainham, D., . . . Osgood, N. (2018). The SMART study, a mobile health and citizen science methodological platform for active living surveillance, integrated knowledge translation, and policy interventions: Longitudinal Study. *JMIR Public Health and Surveillance*, *4*(1), doi: 10.2196/publichealth.8953
- Khoury, M. J., & Ioannidis, J. P. A. (2014). Medicine. Big data meets public health. *Science*, *346*(6213), 1054–1055.
- Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-data Applications in the Government Sector. *Communications of the ACM*, *57*(3), 78–85.
- Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*(1), 1–14.
- Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication and Society*, *20*(1), 14–29.
- Koene, A. (2017). Algorithmic bias: Addressing growing concerns. *IEEE Technology and Society Magazine*, *36*(2), 31–32.
- Kumar, A., Nguyen, V. A., & Teo, K. M. (2016). Commuter

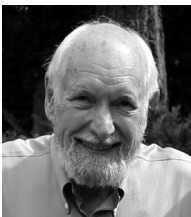
- cycling policy in Singapore: A farecard data analytics based approach. *Annals of Operations Research*, 236(1), 57–73.
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science*, 343(6176), 1203–1205.
- Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., . . . Van Alstyne, M. (2009). Life in the network: the coming age of computational social science. *Science*, 323(5915), 721.
- Lee, D., Kim, M., & Lee, J. (2016). Adoption of green electricity policies: Investigating the role of environmental attitudes via big data-driven search-queries. *Energy Policy*, 90, 187–201.
- Lerner, D., & Lasswell, H. D. (1951). *The policy sciences. Recent developments in scope and method*. Stanford, CA: Stanford University Press.
- Lindblom, C. E. (1959). The science of ‘muddling through’. *Public Administration Review*, 19(2), 79–88.
- Longo, J. (2017). The evolution of citizen and stakeholder engagement in Canada, from Spicer to #hashtags. *Canadian Public Administration*, 60(4), 517–537.
- Longo, J. (2018). Digital tools for rapid policy design. In M. Howlett & I. Mukherjee (Eds.), *Handbook of policy design*. New York, NY: Routledge.
- Longo, J., Kuras, E., Smith, H., Hondula, D. M., & Johnston, E. (2017). Technology use, exposure to natural hazards, and being digitally invisible: Implications for policy analytics. *Policy & Internet*, 9(1), 76–108.
- Longo, J., & McNutt, K. (2018). From policy analysis to policy analytics. In L. Dobuzinskis & M. Howlett (Eds.), *Policy analysis in Canada*. Bristol: Policy Press.
- Lu, W., Chen, X., Ho, D. C. W., & Wang, H. (2016). Analysis of the construction waste management performance in Hong Kong: The public and private sectors compared using big data. *Journal of Cleaner Production*, 112, 521–531.
- Maciejewski, M. (2017). To do more, better, faster and more cheaply: Using big data in public administration. *International Review of Administrative Sciences*, 83(1), 120–135.
- Majone, G. (1989). *Evidence, argument, and persuasion in the policy process*. New Haven, CT: Yale University Press.
- Maragoudakis, M., Loukis, E., & Charalabidis, Y. (2011). A review of opinion mining methods for analyzing citizens’ contributions in public policy debate. In E. Tambouris, A. Macintosh, & H. de Bruijn (Eds.), *Electronic participation (ePart 2011)* (pp. 298–313). Berlin: Springer.
- Masuch, M., & LaPotin, P. (1989). Beyond garbage cans: An AI model of organizational choice. *Administrative Science Quarterly*, 34(1), 38–67.
- May, P. J. (1992). Policy learning and failure. *Journal of Public Policy*, 12(4), 331–354.
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard Business Review*, 90(10), 60–68.
- Meltsner, A. J. (1976). *Policy analysts in the bureaucracy*. Berkeley and Los Angeles: University of California Press.
- Mergel, I., Rethemeyer, R. K., & Isett, K. (2016). Big data in public affairs. *Public Administration Review*, 76(6), 928–937.
- Miller, A. R., & Tucker, C. (2014). Health information exchange, system size and information silos. *Journal of Health Economics*, 33, 28–42.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). doi:10.1177/2053951716679679
- Monafu, D., Chan, S., & Turnbull, S. (2018). Is experimentation just the latest buzzword? *Policy Options*. Retrieved from <http://policyoptions.irpp.org/magazines/february-2018/is-experimentation-just-the-latest-policy-buzzword>
- Morçöl, G. (2001). Positivist beliefs among policy professionals: An empirical investigation. *Policy Sciences*, 34(3/4), 381–401.
- Morgan, M. G., Henrion, M., & Small, M. (1992). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge: Cambridge University Press.
- Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *JAMA: The Journal of the American Medical Association*, 309(13), 1351–1352.
- Murphy, C. (2016). *Shared mobility and the transformation of public transit* (Research Analysis Report #TCRP J-11/TASK 21). Washington, DC: American Public Transportation Association. Retrieved from <https://trid.trb.org/view.aspx?id=1401765>
- Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *IEEE symposium on security and privacy, 2008 (SP 2008)*, pp. 111–125). New York, NY: IEEE.
- Newsham, G. R., & Bowker, B. G. (2010). The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. *Energy Policy*, 38(7), 3289–3296.
- Noveck, B. S. (2018). Legal and other barriers to trust and innovation. In *Workshop on trust and opening governance*. Washington, DC: Georgetown University. Retrieved from <http://opening-governance.org>
- Paris, C., & Wan, S. (2011). Listening to the community: Social media monitoring tasks for improving government services. In *CHI '11 extended abstracts on human factors in computing systems* (pp. 2095–2100). New York, NY: ACM.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.
- Pentland, A. (2012). The new science of building great teams. *Harvard Business Review*, 90(4), 60–69.
- Perry, W. L., McInnis, B., Price, C. C., Smith, S., & Hollywood, J. S. (2013). *Predictive policing: The role*

- of crime forecasting in law enforcement operations. RAND Corporation. doi:10.7249/RR233
- Persily, N. (2017). Can democracy survive the internet? *Journal of Democracy*, 28(2), 63–76.
- Pirog, M. A. (2014). Data will drive innovation in public policy and management research in the next decade. *Journal of Policy Analysis and Management*, 33(2), 537–543.
- Poel, M., Meyer, E. T., & Schroeder, R. (2018). Big data for policymaking: Great expectations, but with limited progress? *Policy & Internet*, 10(3), 347–367.
- Prpić, J., Taihagh, A., & Melton, J. (2015). The fundamentals of policy crowdsourcing. *Policy & Internet*, 7(3), 340–361.
- Quade, E. S. (1975). *Analysis for public decisions*. New York, NY: Elsevier.
- Quade, E. S. (1980). Pitfalls in formulation and modeling. In G. Majone & E. S. Quade (Eds.), *Pitfalls of analysis* (Vol. 8, pp. 23–43). Chichester: John Wiley & Sons.
- Radin, B. A. (2000). *Beyond Machiavelli: Policy analysis comes of age*. Washington, DC: Georgetown University Press.
- Raymer, J., Wiśniowski, A., Forster, J. J., Smith, P. W. F., & Bijak, J. (2013). Integrated modeling of European migration. *Journal of the American Statistical Association*, 108(503), 801–819.
- Robinson, J. B. (1982). Energy backcasting. A proposed method of policy analysis. *Energy Policy*, 10(4), 337–344.
- Russell, S. J., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Essex: Prentice Hall.
- Sawyer, K. R. (2005). *Social emergence: Societies as complex systems*. Cambridge: Cambridge University Press.
- Schintler, L. A., & Kulkarni, R. (2014). Big data for policy analysis: The good, the bad, and the ugly. *The Review of Policy Research*, 31(4), 343–348.
- Shi, J., Ai, X., & Cao, Z. (2017). Can big data improve public policy analysis? In *proceedings of the 18th annual international conference on digital government research* (pp. 552–561). New York, NY: ACM.
- Shulock, N. (1999). The paradox of policy analysis: If it is not used, why do we produce so much of it? *Journal of Policy Analysis and Management*, 18(2), 226–244.
- Sims, C. A. (1986). Are forecasting models usable for policy analysis? *The Quarterly Review*, 10(1), 2–16.
- Stone, D. A. (1988). *Policy paradox and political reason*. Glenview, IL: Scott Foresman & Co.
- Till, B. C., Longo, J., Dobell, A. R., & Driessen, P. F. (2014). Self-organizing maps for latent semantic analysis of free-form text in support of public policy analysis. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, 4(1), 71–86.
- Tingling, P., & Brydon, M. (2010). Is decision-based evidence making necessarily bad? *MIT Sloan Management Review; Cambridge*, 51(4), 71–76.
- Tsoukias, A., Montibeller, G., Lucertini, G., & Belton, V. (2013). Policy analytics: An agenda for research and practice. *EURO Journal on Decision Processes*, 1(1/2), 115–134.
- Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday*, 19(7). doi:10.5210/fm.v19i7.4901
- Van Oort, N., & Cats, O. (2015). Improving public transport decision making, planning and operations by using big data: Cases from Sweden and the Netherlands. In *18th IEEE international conference on intelligent transportation systems, Las Palmas, Spain, 15–18 September 2015*. New York, NY: IEEE.
- Verhulst, S. G., & Young, A. (2018). The potential and practice of data collaboratives for migration. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/the_potential_and_practice_of_data_collaboratives_for_migration
- Vie, L. L., Griffith, K. N., Scheier, L. M., Lester, P. B., & Seligman, M. E. P. (2013). The person-event data environment: Leveraging big data for studies of psychological strengths in soldiers. *Frontiers in Psychology*, 4, 934.
- Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126, 3–13.
- Ward, J. S., & Barker, A. (2013). Undefined by data: A survey of big data definitions. *arXiv [cs.DB]*. Retrieved from <http://arxiv.org/abs/1309.5821>
- Welch, E. W., Hinnant, C. C., & Moon, M. J. (2005). Linking citizen satisfaction with e-government and trust in government. *Journal of Public Administration Research and Theory*, 15(3), 371–391.
- Wigan, M. R., & Clarke, R. (2013). Big data's big unintended consequences. *Computer*, 46(6), 46–53.
- Wildavsky, A. (1979). Policy analysis is what information systems are not. In A. Wildavsky (Ed.), *The art and craft of policy analysis* (pp. 26–40). London: Palgrave Macmillan.
- Xiong, G., Zhu, F., Dong, X., Fan, H., Hu, B., Kong, Q., . . . Teng, T. (2016). A kind of novel ITS based on space-air-ground big-data. *IEEE Intelligent Transportation Systems Magazine*, 8(1), 10–22.
- Zhou, K., & Yang, S. (2016). Understanding household energy consumption behavior: The contribution of energy big data analytics. *Renewable and Sustainable Energy Reviews*, 56, 810–819.

About the Authors



Justin Longo is the Cisco Research Chair in Digital Governance and an Assistant Professor in the Johnson-Shoyama Graduate School of Public Policy at the University of Regina where he directs the Digital Governance Lab. He has a PhD in public policy and public administration from the University of Victoria (2013) and was a Postdoctoral Fellow in open governance at Arizona State University. His current research focuses on the organizational and societal implications of advancing technology.



Rod Dobell is Emeritus Professor of Public Policy at the University of Victoria, and Senior Research Associate of the Centre for Global Studies and the Centre for Public Sector Studies. He completed his PhD in economics at the Massachusetts Institute of Technology, and taught at Harvard University, the University of Toronto, and the University of Victoria. In 1991 he was named as the first appointee to the Francis G. Winspear Chair for Research in Public Policy at the University of Victoria.