Government Information Quarterly 33 (2016) 371–377

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

The challenges and limits of big data algorithms in technocratic governance

Marijn Janssen [a,*], George Kuk [b]

[a] Delft University of Technology, The Netherlands
[b] Nottingham Trent University, UK

ARTICLE INFO

Keywords:
big data
algorithm
governmentality
materiality
computational governance
technocratic governance
curation
transparency
accountability
smart cities
cybersecurity

ABSTRACT

Big data is driving the use of algorithm in governing mundane but mission-critical tasks. Algorithms seldom operate on their own and their (dis)utilities are dependent on the everyday aspects of data capture, processing and utilization. *However, as algorithms become increasingly autonomous and invisible, they become harder for the public to detect and scrutinize their impartiality status*. Algorithms can systematically introduce inadvertent bias, reinforce historical discrimination, favor a political orientation or reinforce undesired practices. Yet it is difficult to hold algorithms accountable as they continuously evolve with technologies, systems, data and people, the ebb and flow of policy priorities, and the clashes between new and old institutional logics. Greater openness and transparency do not necessarily improve understanding. In this editorial we argue that through unravelling the imperceptibility, materiality and governmentality of how *algorithms work*, we can better tackle the inherent challenges in the curatorial practice of data and algorithm. Fruitful avenues for further research on using algorithm to harness the merits and utilities of a computational form of technocratic governance are presented.

1. Introduction

The use of algorithm is moving to the heart of the governance of our society as more and more is digitized. Important decisions about people are increasingly made by algorithm (Kroll, 2015). In recent years, we have witnessed the rise of social algorithms, which "*size us up, evaluate what we want, and provide a customized experience*" (Lazer, 2015) and an increase use of algorithm by governments to systemize and scale down bureaucracy and improve their decision-making. The rise of algorithm is driven by the availability of Big and Open Linked Data (BOLD) (Janssen & Kuk, 2016). The implications of algorithm might vary and can be either positive or negative, for example by decreasing privacy and by increasing transparency (Janssen & Van den Hoven, 2015). Not big data, but the use of algorithm to process data might be the real hidden danger.

Typical examples of the use of algorithm can already be found in the fields of searching, surveillance, traffic management, decision-making, fraud and smart cities. Algorithms can help to find information. They can be used to identify persons in pictures, detect the occurring of violence in video surveillance, and detect anomalies and unusual patterns in tax return data for fraud identification. They can be also used to control and regulate traffic flows suggesting the speed and lanes. Also within the government operational processes algorithms are used, for example to process and analyze crime reports (Ku & Leroy, 2014) and to make decisions whether to grant permits. Another example of use is to calculate the risks of re-offending of criminals (Eveleth, 2016). Often we are unaware of the use of algorithm in our daily life, but slowly they are invading our society.

Generally speaking, algorithm consists of step-by-step processes and/or rules processing inputs into outputs (Stone, 1972). Whereas algorithms were primarily viewed as the domain of computer programmers in the past, algorithms have become central to the ways societies and economies are run and governed. References to "the Google algorithm" and so on are typically shorthand for high-level assemblies of many sub-algorithms, all eventually implemented and running as computer programs (Sandvig, 2014). The idea of algorithm refers to some intelligence that manipulates data resulting in an outcome. Algorithm can often not be depicted by a simple flowchart or other representational mechanism, and algorithm might not be able to be separated from its implementation (Sandvig, 2014). Nonetheless understanding algorithms does require in-depth knowledge of mathematics and even computer science to comprehend how they work.

Using algorithms is central to the computational approach in the design and operations of our society. Algorithms governing our society results in *technocratic governance*. This concept builds on the premise that "all aspects of a city can be measured and monitored and treated as technical problems which can be addressed through technical solutions" (Kitchin, 2014, p. 9). Technocratic governance assumes that complex societal problems can be deconstructed into neatly defined,

* Corresponding author.
E-mail address: M.F.W.H.A.Janssen@tudelft.nl (M. Janssen).

http://dx.doi.org/10.1016/j.giq.2016.08.011

structured and well-scoped problems that can be solved algorithmically and in which political realities play no role. The political and social solutions play a lesser role in that both the governing and the governed are subjugated to the control of technology. This presents a new form of technological determinism. Technocratic governance is often found in the beliefs that neutrality can be attained by removing humans and relying upon algorithmic automation.

This view is in contrast with recent studies showing that the use of algorithm can systematically introduce inadvertent bias (e.g. Kashin, King, & Soneji, 2015), reinforce historical discrimination (Kroll, 2015), filter out information (Pariser, 2011), or reinforce outdated practices or failures from the past, whereas societal values and political preferences have changed. Often data mining simply unearths prior biases and discriminatory practices, and reveals the entrenched social inequalities (Barocas & Selbst, 2016). Furthermore, algorithms can be used to nudge users and influence the motives, behavior and decision of individuals and groups (Thaler & Sunstein, 2008).

Both the design and use implications of algorithm need to be unraveled to create transparency and accountability. Fig. 1 displays a simplistic overview starting with the collection of data using all kinds of devices resulting in the collection of data followed by an algorithm used for processing the data and ending with use. This constitutes a simplified version of our black-box understanding of algorithm. When taking other aspects of algorithm into account, including their inception, design, creation, implementation and use, it becomes clear that this process is much more complex. In this editorial, we will discuss some of the use contexts of algorithm to underline its complexity and materiality. The latter has no fixed boundaries. Algorithms dynamically co-evolves with data, systems and humans within complex socio-technical system.

The issues of accountability remain and prod a contentious discussion about the power and control of algorithm in the inclusion and exclusion of information, and the consequences and disutilities when they fail to perform and unintentionally introduce systematic biases in our decision making (e.g. Kashin et al., 2015). Some of these systematic biases are inherited with the historical databases that algorithm use, whereas others are unintentionally introduced by designers or originate from changes in the systems and data landscape. In effect, *decisions produced by the algorithms are as good as the data upon which such decisions are computed and the humans and systems operating them*.

Given its impact on society there is a need for openness and transparency for the design, operating and impact of using algorithm to govern our society Several questions remain when it comes to the curatorial practices of data and algorithm. What are the checks and balances for algorithmic control and behaviours? Can citizens scrutinize algorithms especially where algorithms are closely guarded and require technical know-how? The question is if the hidden details of how algorithms work become more transparent, can we engage in a more effective critique and in turn make the algorithms more accountable? Or will the transparency simply render the algorithm obsolete by making them readily exploitable by malicious actors in gaming the systems and inducing more disutility to the citizens and negating the transparency benefits?

2. Growing complexity and transparency

Algorithms can perform simple calculations or highly complex reasoning tasks. Algorithms can be processed by humans and also by computer programs. Although algorithm has been around for a long time, nowadays algorithm are increasingly becoming an integral part of computing programs. In this way algorithm become increasingly autonomous and invisible invading the governance of daily activities. Algorithm can be expressed in many kinds of notation, including natural, arithmetic or computer languages. For execution, algorithm are often implemented in programming languages to allow computers to process them., Algorithm should be independent of the mechanisms implementing them, however, this neglects the complexity of representing algorithm (Knuth, 1966). Some algorithms are too complex, in that they cannot be separated from their implementation anymore.

Fig. 2 shows the type of algorithm can be categorized based on their level of automation and how complex they are. Contemporary examples of algorithm used for decisions-making and governance include the following.

- An example of a *simple* algorithm that can be *manually* operated is the checking of your identity by civil servants. Various authenticity features are checked and the picture on the identity card is compared with the person stating in front of the civil servants. At customs this process becomes more and more automated.
- An example of a *simple* and often *automated* algorithm is calculation of social benefits. Legislations and policy-makers provides the conditions under which a person is eligible and based on the situation a certain fee is applied. The characteristics of the person and the situations are used as an input and the algorithm calculates the resulting social benefits.
- An example of a *complex and manual* operated algorithm is the admissions of immigrants. Often many trade-offs and considerations are taking into account. They are guided by multiple legislations that require interpretation by policy-makers and lawyers. Due to the many variables and situations it is difficult to automate fully. It will rely on the discretion of the officer(s) in charge to decide on a case-by-case basis.
- A *complex and automated* algorithm is the sequence of activities used to determine if passengers might be a security threat. Data stored in database from multiple security agencies are combined with the characteristics of the travelers to determine if there is a possible security risk. Based on these outcomes, actions can be taken.

In each of the examples given there is a tendency to move to higher level of automation and complexity as visualized in Fig. 2. In the current society more and more algorithm are embedded in computer programs. This enables us to deal with more complex situations than any single persons can handle, at the expense of easiness to understand its functioning. Algorithm grow in complexity by combining many algorithms

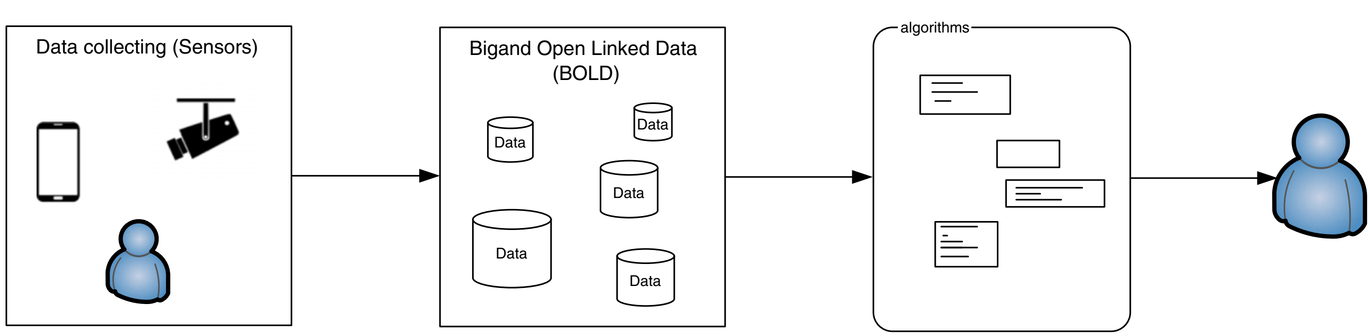

Fig. 1. Simple conceptualization of data algorithm and use