# Algorithms and Public Policy


| | |
|---|---|
| **Name** | **J. R. Kroeskop** |
| **Student number** | **3412261** |
| **Supervisor** | **Prof Dr. F. Brom** |
| **Second reader** | **Dr. F.L.B. Meijboom** |
| **Date** | **22-06-2018** |

**Abstract**

In this paper, I will discuss what the ethically relevant aspects are for employing algorithms in the public sector. To do this, I will first show what algorithms are, and that they should be seen as more than mere mathematical constructs, on the grounds of that citizens experience them as part of wider systems that influences their lives. Afterwards, I will give five interlocking arguments for and against the use of algorithms in the public sector. These arguments will be combined with the welfare utilitarianism propagated by Robert Goodin to develop an ethical framework for the normative evaluation of the use of algorithms in the public sector.[1] In the final section, I will consider the merits and demerits of predictive policing on the basis of the ethical framework.

[1] Robert E. Goodin, *Utilitarianism as a public philosophy*. (New York: Cambridge University Press, 199), 23.

**Research question**

What are the ethically relevant aspects of the employment of algorithms in the public sector with regards to maximizing public welfare?

**Introduction**

Applying for a loan has become a significantly different process in the age of Big Data. This difference mostly consists of how the decision is being made whether or not someone qualifies for the loan. Frank Pasquale, a legal scholar that is specialized in the influence algorithms, describes how this assessment has mostly become automated.[2] Before this automation, the decision whether someone qualifies would be made by a bank clerk. The clerk would base his or her decision on the financial history of the applicant, presumably ask questions about anything the blank clerk finds concerning, and try to make a judgment partly based on objective criteria, and partly based on experience. This judgment could be wrong or even influenced by human bias. In corner cases, the bank clerk would have to rely on his or her gut feeling about the applicant. This leaves an uncomfortable amount of room for prejudice to creep in.

Automating the decision-making process involves employing algorithms to sort through the relevant data to come up with some kind of score that is correct in a sufficient amount of the time. For instance, an algorithm might sort through the data about the loan applicant, compare that data with the data on applicants that proved to be able to pay back the loan and come up with a score that would signal the creditworthiness of the loan applicant. The main arguments for automating such tasks is that is it leads to better results, due to the algorithms making fewer errors than their human counterparts. It is also more efficient because algorithms require less time and recourses to perform the task of data comparison. Cathy O'Neil, a mathematician and data-scientist, also states that the introduction of FICO-scores, an algorithm that only considers an applicant's financial history as opposed to other factors such as zip code or race, has made the application system more objective and fair.[3] However, she argues that modern e-scoring, a system where large amounts of data about a person are being processed by algorithms to generate an image of that person, re-introduces all the factors that the FICO-score has left out. She gives as an example how a website of a credit card company might process all kinds of information, such as at what neighborhood a person is from or what kind of car that person drives, in order to raise the interest rate for that person or limit the amount of credit that person is being offered.[4]

---

[2] Frank Pasquale, *The Black Box Society.* (Cambridge, Massachusetts: Harvard University Press, 2016), 22-25.
[3] Cathy O'Neil. *Weapons Of Math Destruction.*( London: Penguin Books, 2011), 142.
[4] O'Neil, *Weapons of Math Destruction*, 142, 143.

Pasquale is highly critical of the FICO-scores. He argues that these scores introduce arbitrary rules that are difficult to make sense of.[5] Pasquale gives several examples of such rules. For instance, the score considers what proportion of credit is being used. This means that if a person would decide to set a lower limit to his or her credit card, while having legitimate reasons for doing so, and not borrowing any extra money, that person would significantly lower his or her credit rating.[6] Moreover, it punishes those that have not taken on any loans in the past, due to there being no data on their credit history.[7] The negative consequences of having a low credit-score can be severe. They include not having access to credit or only at high interest rates. Pasquale argues that the negative consequences go even further because credit-scores are being used for many other purposes. For instance, HR managers use them as an indication of reliability when making hiring decisions.[8] Pasquale also questions whether or not these credit systems are truly unbiased. He claims that if due to human bias, minorities historically have received loans at higher interest rates, and having higher interest relates correlates with a higher chance to default on that loan, the data would ultimately reinforce the human bias.[9] Therefore, even if the system appears to be entirely objective, and all of the present steps are free from human bias, the algorithm can still inherit the bias from the historic data set.

The FICO-score example illustrates how normative research to the consequences of the employment of algorithms is becoming increasingly relevant due to the far-reaching consequences it has on the daily lives of individuals. The complications a person can be faced with on the basis of a low credit-score will go as far as having difficulties getting a job, buying a house or even getting a phone subscription. Such complications become a far greater concern when coupled with public policy. The potential influence public institutions can have on the lives of an individual are not only far greater, they are also difficult to escape from. This is not to say public institutions ought to refrain from employing algorithms or pursuing automation. The benefits provided by the use of algorithms are simply too great to leave on the table. It necessitates taking into consideration the merits and demerits of employing an algorithm, and assessing these in a structured fashion. To do this a framework is needed to structure the prevalent arguments for and against the use of algorithms. This framework should draw from a wide array of fields in order for it to be relevant for a broad set of real-world cases.

My purpose in this paper will be to give an outline of such an framework. To do this, I shall discuss the ethically relevant aspects of employing algorithms by going into the arguments for and against the use of algorithms. These interlocking arguments will then be combined with the public welfare utilitarianism propagated by Robert Goodin.[10] He argues that public

[5] Pasquale, *The Black Box Society*, 22-25, 41.
[6] Pasquale, *The Black Box Society*, 23.
[7] Pasquale, *The Black Box Society*, 24.
[8] Pasquale, *The Black Box Society*, 25.
[9] Pasquale, *The Black Box Society*, 41.
[10] Goodin, *Utilitarianism as a public philosophy*, 23.

servants are obligated to act in such a way that it maximizes the public interest. Such a position is well suited for assessing the influence of the use of algorithms in the public sector because it provides a normative component to the arguments for and against the use of algorithms. For instance, for public policy-makers efficiency can be seen as a normative duty because it increases the services that the public sector is able to provide. Goodin claims that such a duty can be derived from the responsibilities that are associated with public service.[11] Based on these responsibilities of public policy-makers it is relevant to consider how the implementation of a policy relates to maximizing public interest. In the final section I intend to do this by applying the ethical framework to a specific case study.

In the first section, I will provide an overview of the debate about what algorithms are and how they influence our lives, in order to identify what should be looked at when researching the employment of algorithms in the public sector. To do this I will consider the distinction made by Robin Hill between algorithms as mathematical constructs and algorithms as they are perceived by the public, and argue that the latter is the most pertinent for an ethical framework.[12] I will base this on the notion of the performative nature of algorithms discussed by Rob Kitchin. How citizens experience, and how they are affected by, algorithms is bound to be as constituted within a wider system. Therefore, this entire system should be considered.[13]

In the second section, I will discuss the five main arguments for the use of algorithms. The first of these is that the employment of algorithms enables new kinds of opportunities for public policy-makers. An example that illustrates this is that, according to Achrekar et al., public policy-makers can use social media data to take preemptive actions during influenza outbreaks.[14] The second argument for the employment of algorithms is that the can be used to increase efficiency. For instance, Rhoda Joseph and Johnson, argue that public institutions can become more efficient by using algorithms to automate a large number of transactions made with citizens.[15] The third and fourth arguments address how the use of algorithms will make the way organizations interact with the public more reliable. This can be done by reducing the influence of human error or human bias. An example of this is that, according to Gauri Naik and Sanika Bhide, the implementation of medical diagnosis software will cause medical professionals to make better decisions with regards to the health of their patients.[16] The fifth argument is that algorithms are free from human bias. From the discussion about

[11] Robert, E Goodin. “Public Service Utilitarianism as a Role Responsibility.” *Utilitas*, 10(3), (1998): 322.
[12] Robin K. Hill,"What an algorithm is." *Philosophy & Technology*29, no. 1, (2016): 36.
[13] Rob Kitchin, "Thinking critically about and researching algorithms." *Information, Communication & Society* 20, no. 1 (2017): 16.
[14] Harshavardhan Achrekar, Gandhe Avinash, Ross Lazarus, Ssu-Hsin Yu, and Benyuan Liu, "Predicting flu trends using twitter data." In *Computer Communications Workshops* (2011): 716, 717, 718.
[15] Rhoda C. Joseph and Norman A. Johnson, "Big data and transformational government." *IT Professional* 15, no. 6 (2013): 46.
[16] Gauri Naik, and Bhide S. Sanika, "Will the future of knowledge work automation transform personalized medicine?" *Applied & translational genomics* 3, no. 3 (2014): 50-53.

FICO-scores by O'Neil it is clear that algorithmic scoring systems can be used to reduce the influence of human bias during the decision-making process.[17]

In the third section, I will give a counterargument to each of the arguments for the use of algorithms. The first of these is about how human bias can creep into algorithms. For instance, according to Kraemer et al., in some cases the development of an algorithm requires taking a stand in an ethical debate because the designer has to make a choice about how to manage tradeoffs.[18] Secondly, I will consider that there are necessary limits to how efficient an organization should. Casey Haskins even argues that a small margin of inefficiency is vital for a public institution to function properly.[19] The third argument addresses the complications that arise once an algorithm becomes ubiquitous. For instance, O'Neil claims that the use of hiring algorithms has proliferated to such a degree that certain groups of people will find it significantly more difficult to find a job.[20] The fourth argument addresses the peculiar mistakes made by algorithms. Fifthly, I will discuss the importance of maintaining a role for human decision-makers. Brenninkmeijer argues that citizens should have the ability to appeal to a human decision-maker. This human decision-maker would necessarily have to have the discretionary power to bend the rules of an algorithmic system, thereby, building in a limitation to which degree a system can be automated.[21]

In the fourth section, I will combine these five interlocking arguments with the public welfare utilitarianism proposed by Goodin to develop an ethical framework to assess the use of an algorithmic system. For a discussion about whether the implementation of an algorithm adds or subtracts to the maximizing of public welfare a distinct advantage is the egalitarian bend described by Goodin.[22] Poignantly, this egalitarian bend needs to be emphasized when assessing the use of an algorithm in the public sector on the grounds of the severe negative consequences it can have on marginalized groups.

In the fifth section, I will apply the framework to the case study of predictive policing. I will discuss predictive policing in the light of two legs of the ethical framework. The first will be about how predictive policing fair with regards to maximizing public welfare through gains in efficiency, and how this might be limited on the grounds of necessary inefficiencies. Secondly, I will go into the notion of the objectivity of the algorithms making the predictions, and how this relates to the possibility of biased outcomes.

---

[17] O'Neil, *Weapons of Math Destruction*, 142.
[18] Felicitas Kraemer, Kees Van Overveld, and Martin Peterson, "Is there an ethics of algorithms?" *Ethics and Information Technology* 13, no. 3 (2011): 258.
[19] Casey P Haskins, "The drawbacks of efficiency?" *The Chautauquan Daily*, August 16, 2011. https://chqdaily.wordpress.com/2011/08/16/the-drawbacks-of-efficiency/.
[20] O'Neil, *Weapons of Math Destruction*, 100.
[21] Alex Brenninkmeijer, "Meer dan een burger-servicenummer." Accessed June 20, 2018. https://www.nationaleombudsman.nl/uploads/tijdschrift_voor_conflicthantering_hoeveel_legitimatie_schept_lokale_democratie.pdf
[22] Goodin, *Utilitarianism as a public philosophy*, 23.

# 1. Algorithmic Systems

## Introduction

In this section, I intend to give an overview of what algorithms are and when they become relevant for normative research. Afterwards, I will show how they relate to Big Data and data analysis. Finally, I will argue that algorithms should be seen within the scope of the systems they are involved in. According to Rob Kitchin, algorithms are performative in nature.[23] Consequently, algorithms cannot be seen separately from the tasks they perform or enable to be performed. Added to this is that they should be regarded in light of how they influence our lives and the systems they invoke. This means that algorithms should be understood as algorithmic systems that involve aspects such as how users interact with them, the design choices that are made during their development, the databases that are fed to them, and how they are likely to function in the future.[24]

## 1.1 Defining algorithms

To study how the employment of algorithms influences our lives it is important to look into several definitions of algorithms provided by different authors in order to identify the most relevant aspects of what algorithms are and do. These aspects will provide the basis for a list of desiderata for a definition of the kind of algorithms that are salient for normative study.

The first is the definition used by the programmers and engineers that create those algorithms. In *Introduction to the Design and Analysis of Algorithms* they are defined as:

“An algorithm is a sequence of unambiguous instructions for solving a problem, i.e., for obtaining a required output for any legitimate input in a finite amount of time.” [25]

The key component of this definition is that an algorithm can be considered as a sequence of steps, here called instructions, that produce a result, namely the solving of a problem. Kitchin gives a similar, yet more succinct, definition. Kitchin defines algorithms as:

“sets of defined steps structured to process instructions/data to produce an output.”[26]

Noteworthy is that Kitchin places an emphasis on the relation between the input that goes into an algorithm, and the output generated by the algorithm. This is problematic because it remains vague about what a valid output might be, and how this output is determined. Furthermore, it ties the data and instructions that provide the input of the algorithm together, while leaving unspecified what exactly the role of both these aspects is. To study

---

[23] Kitchin “Thinking critically about and researching algorithms” 16, 25.
[24] Kitchin “Thinking critically about and researching algorithms” 25.
[25] Anany Levitin, *Introduction to the design & analysis of algorithms*. (Boston: Pearson, 2012), 3.
[26] Kitchin “Thinking critically about and researching algorithms” 14.

how the employment of algorithms influence our lives it is necessary to make a clear distinction between those aspects, in order to identify what consequence is caused by what particular design choice or bias in a data set. Answers to these questions might be found with a more elaborate definition proposed by Tarleton Gillespie, considering the definition given by Kitchin is based on that of Gillespie. According to Gillespie, an algorithm is:

"Algorithms need not be software: in the broadest sense, they are encoded procedures for transforming input data into a desired output, based on specified calculations. The procedures name both a problem and the steps by which it should be solved."[27]

Gillespie makes a distinction between the input, the data set, and the output or what Kitchin called the solving of a problem. This enables Gillespie to leave room in his definition for the choices that need to be made by either the designer of an algorithm or the client during the development of an algorithm. Achieving a desired output requires the client to specify what would qualify as a desired output.[28] It are exactly these choices that open the door for human bias to step through. It also shows what the algorithm looks like, and how it will function, is to some degree dependent on the choices made by both the client and the developer.

What is lacking in Gillespie's definition are the formal requirements to the set of instructions and the input. In the definition provided by *Introduction to the Design and Analysis of Algorithms*, there was an emphasis on the unambiguity and the validity of the input. Both these requirements are quintessential to deliver a functioning algorithm. In Gillespie's definition, they might very well be assumed. However, the requirement of unambiguity needs to be in a definition of an algorithm, especially since ambiguity in the set of instructions will lead to errors and unpredictable outcomes.[29]

Kraemer et al.[30] give a definition that incorporates these requirements. Their definition also avoids the problem described by Hill. Namely that strictly formal definition would only be confusing to the public, rather than adding to the understanding that is needed to allow them to partake in a meaningful debate about the ethical concerns regarding the use of algorithms.[31] The definition proposed by Kraemer et al. is:

"An algorithm is, roughly speaking, a finite sequence of well-defined instructions that describe in sufficiently great detail how to solve a problem."[32]

This definition is less succinct than that proposed by Kitchin.[33] However, it is still relatively vague when compared to the definition found in *Introduction to the Design and Analysis of*

---

[27] Tarleton Gillespie, "The relevance of algorithms." *Media technologies: Essays on communication, materiality, and society*, (2014): 1.
[28] Gillespie, "The relevance of algorithms" 1.
[29] Levitin, "Introduction to the design & analysis of algorithms," 3.
[30] Felicitas Kraemer, Kees van Overveld, Martin Peterson
[31] Robin K Hill, "What an algorithm is." *Philosophy & Technology*29, no. 1, (2016): 37.
[32] Kraemer, et al., "Is there an ethics of algorithms?" 251.
[33] Hill, "What an algorithm is" 37

*Algorithms*. It also lacks the highly relevant aspects of input, especially data sets, and subjective design choices mentioned in the definition proposed by Gillespie. Both these aspects are crucial for a normative argument regarding the consequences of the employment of algorithms.

John Danaher gives a definition that places an emphasis on the relation between input and output. Furthermore, his definition is notably different from that of Gillespie due to Danaher limiting algorithms to the realm of computers. Danaher's definition is:

"algorithms, i.e. computer programmed step-by-step instructions for taking a given set of inputs and producing an output."[34]

The definition suggested by Danaher faces the same problem as that proposed by Kitchin. Although it is elegant and manages to places an emphasis on the relation between input and output, it is too sparse to address the relevant ethical concerns. A definition needs to at least contain more information on what kind of input is used, and it needs to address what kind of output is generated.

Mittelstadt et al. do not give a clear a definition of what an algorithm is. They prefer the formal definition proposed by Hill for algorithms as mathematical constructs.[35] However, as Mittelstadt et al. rightfully point out, for ethical research the interesting algorithms are those that perform tasks, especially those that make decisions that are difficult for humans to comprehend.[36] To solidify this point it is useful to look into the difference between the two formal definitions proposed by Hill. According to Mittelstadt et al., the "popular usage becomes relevant"[37] precisely with this difference.[38]

## 1.2 Formal definition for algorithms as mathematical objects

The first definition proposed by Hill is:

"An algorithm is a finite, abstract, effective, compound control structure, imperatively given."[39]

This concise formal definition has several interesting aspects. First of all, the inclusion of an algorithm being effective entails that it needs to "be certain to produce the result"[40]. This excludes, according to Hill, cooking recipes from being algorithms. Hill claims that cooking

[34] Danaher, John. "The threat of algocracy: Reality, resistance and accommodation." *Philosophy & Technology* 29, no. 3 (2016): 251.
[35] Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. "The ethics of algorithms: Mapping the debate." *Big Data & Society* 3, no. 2 (2016): 2.
[36] Mittelstadt et al., "The ethics of algorithms: Mapping the debate" 3.
[37] Mittelstadt et al., "The ethics of algorithms: Mapping the debate" 2.
[38] Mittelstadt et al., "The ethics of algorithms: Mapping the debate" 2, 3.
[39] Hill, "What an algorithm is" 44.
[40] Hill, "What an algorithm is" 45.

recipes are a common example of what an algorithm is.[41] However, Hill argues that recipes would not meet the requirements of the definition proposed by her due to the possibility of the instructions being carried out badly. This possibility would mean that the instructions might be carried out in such a haphazard manner that it cannot be guaranteed that the desired result will be achieved.[42]

An interesting aspect of Hill's first definition is that it does not consider the "human point of view"[43], which Hill admits needs to be admitted to provide a comprehensive definition of what an algorithm is.[44] Therefore, introduces an amendment to her definition:

"An algorithm is a finite, abstract, effective, compound control structure, imperatively given, accomplishing a given purpose under given provisions."[45]

Mittelstadt et al. identify that the added criterion of "accomplishing a given purpose under given provisions"[46] is where algorithms become interesting for normative investigation.[47] They argue that: "References to algorithms in public discourse do not normally address algorithms as mathematical constructs, but rather particular implementations. Lay usage of 'algorithm' also includes *implementation* of the mathematical construct into a technology and an application of the technology *configured* for a particular task."[48] According to Mittelstadt et al., it makes sense for the public discourse to be about how the algorithms are implemented, and the tasks they perform, because that is when they start affecting our daily lives.[49]

### 1.3 Defining ethically relevant algorithms

From the definitions proposed by Kitchin, Gillespie, Danaher, and Kraemer et al. it is possible to give a list of desiderata for what would be required for a definition that addresses those algorithms that are most influential in shaping daily life. Such a list of desiderata would focus on those aspects of an algorithm that allow it to be consequential for our daily lives.

Based on the definitions discussed in the previous sections it is possible to give a list of desiderata for a definition of algorithms.

- ❖ An algorithm can be considered ethically relevant when it impacts our daily lives or is not necessarily precluded from doing so.

---

[41] Hill, "What an algorithm is" 49.
[42] Hill, "What an algorithm is "49.
[43] Hill, "What an algorithm is" 45.
[44] Hill, What an algorithm is" 46.
[45] Hill, "What an algorithm is" 47.
[46] Hill, "What an algorithm is" 47.
[47] Mittelstadt et al., "The ethics of algorithms: Mapping the debate" 2.
[48] Mittelstadt et al., "The ethics of algorithms: Mapping the debate" 2.
[49] Mittelstadt et al., "The ethics of algorithms: Mapping the debate" 2, 3.

- A definition must explain that an algorithm contains a sequence or set of instructions, that are somehow confined.
- A definition must clearly state that an algorithm produces an outcome or output.
- A definition must lay out the relation between this output and the input the algorithm uses to produce that output.
- A definition must state what this input might consist of, and make a clear separation between this input and the set of instructions.
- A definition must leave room for the subjective influence of the designer or client.

Based on this list, and the definition proposed by Gillespie,[50] a definition for what an algorithm is can be formulated. This would be:

*An algorithm is a procedure that involves a confined sequence of unambiguous instructions in order to deliver an output, a specified and previously determined outcome, from a valid input, a data set. This output is capable of affecting daily human life, and has been imagined either during the design process of the algorithm or while the algorithm is being employed.*

### 1.4 Big Data and Algorithms

Algorithms play a vital role in enabling the progress made through the rise of Big Data. Boyd and Crawford, both scholars working for Microsoft Research, observed that: “Big Data is less about data that is big than it is about a capacity to search, aggregate, and cross-reference large data sets.”.[51] This capacity is for a large part provided by algorithms. Algorithms do the work that is required for Big Data practices to thrive. This is because algorithms are crucial for the process of data mining, discovering patterns in data sets,[52] and what Höchtl et al.[53] call Big Data Analytics, the process of extracting meaningful insights from large data sets.[54] Höchtl et al., researchers specialized in public policy, also argue that the kind of algorithms used for Big Data Analysis is required to meet specific demands in scalability, being able to process ever-larger data sets, timeliness, the ability to do this in real-time, and organization, who gets access to what information at which time.[55] Chris Yiu, a public policy researcher, adds to this that public officials will have to employ machine learning algorithms to gain insights through Big Data analysis. He claims it is impossible to gain these insights when merely relying on traditional data analysis techniques and human mental capabilities.[56] Yiu

---

[50] Gillespie, “The relevance of algorithms“ 1.
[51] Danah Boyd, and Kate Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon." *Information, communication & society* 15, no. 5, (2012): 663.
[52] David J. Hand, "Principles of data mining." *Drug safety* 30, no. 7 (2007): 62.
[53] Johann Höchtl, Peter Parycek, and Ralph Schöllhammer
[54] Höchtl, Johann, Peter Parycek, and Ralph Schöllhammer. "Big data in the policy cycle: Policy decision making in the digital era." *Journal of Organizational Computing and Electronic Commerce* 26, no. 1-2 (2016): 153,154.
[55] Johann Höchtl, Peter Parycek, and Ralph Schöllhammer, "Big data in the policy cycle: Policy decision making in the digital era." *Journal of Organizational Computing and Electronic Commerce* 26, no. 1-2 (2016): 154.
[56] Chris Yiu, "The big data opportunity." *Policy exchange* 8 (2012), 10, 11, 15.

argues that this is because Big Data is essentially about data sets have become sufficiently complex and unwieldy that they go beyond what human understanding, combined with the hands-on analysis is capable, to process.[57]

Höchtl et al. state that the kind of data that is being collected also encourages the employment of machine learning algorithms.[58] Kim et al.[59] observe in *Big-Data Applications in the Government Sector* that the vast majority, close to 90%, of data that is being collected, is unstructured data.[60] According to Höchtl et al., this entails that: "for a computer system the effort to automatically derive meaningful insights is much higher than in the case of structured data"[61]. They further argue that the volume of unstructured data that needs to be processed brings it beyond what humans would manually be able to achieve.[62] This will only intensify given that, according to Kim et al., the amount of unstructured data that is being collected is growing at an amazing rate, with around 2.25 quintillion bytes or 2 250 000 Terabytes per day.[63]

## 1.5 Choosing the Algorithm

It remains underappreciated that the choice of what kind of algorithm is to be used for this process also has a significant impact on what results will be achieved. A panel that studied the most important data mining algorithms claim that certain algorithms can produce results more efficiently, however, these are also prone to making specific kinds of errors while other kinds of algorithms are capable to sift through exceedingly complex data sets at the cost of human intelligibility.[64] This means that the choice of what kind of algorithm is to be employed has normative salience. Algorithms are not just dormant, value-neutral, objects. When one is chosen over the other it brings with it consequences for the subjects of the tasks performed by these algorithms. This necessitates a debate about the ethics of these choices.

57 Yiu, "The Big Data Opportunity," 10, 11.
58 Höchtl, Johann, Peter Parycek, and Ralph Schöllhammer. "Big data in the policy cycle: Policy decision making in the digital era." *Journal of Organizational Computing and Electronic Commerce* 26, no. 1-2 (2016): 153.
59 Gang-Hoon Kim, Silvana Trimi, and Ji-Hyong Chung
60 Gang-Hoon Kim, Silvana Trimi, and Ji-Hyong Chung. "Big-data applications in the government sector." *Communications of the ACM* 57, no. 3 (2014): 78.
61 Höchtl, et al., "Big data in the policy cycle: Policy decision making in the digital era" 153.
62 Höchtl, et al., "Big data in the policy cycle: Policy decision making in the digital era" 153.
63 Jacobson, Ralph. "2.5 quintillion bytes of data created every day. How does CPG & Retail manage it?" *IBM*, April 24, 2013. https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/
64 Xindong Wu, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan et al, "Top 10 algorithms in data mining." *Knowledge and information systems* 14, no. 1. 4, 5, 12, 30.

## 1.6 Algorithmic systems

From the example of loan applications, it is clear that the employment of algorithms has real-world implications for individuals. A person that would not have been able to get a loan before the introduction of FISCO-scores might now be able to get access to a line of credit. Contrariwise, a person that would have qualified before might now be turned down or would have to pay a substantially higher interest rate. Personalized marketing can have the adverse effect of allowing landlords to only advertise to specific demographic groups or those that fit pre-determined criteria. Kate Crawford and Jason Schult warn that such use of algorithms would enable landlords to circumvent anti-discrimination laws. The employment of algorithms would allow landlords to target specific groups to show their advertisements to, thereby, denying members of other groups the opportunity to respond to the advertisement. Crawford and Schultz warn that members of certain groups will be entirely left out, because they will not be able to see the advertisements.[65] ZipRecruiter, an online job board, uses algorithms to find out what kind of candidate a company was previously interested in, and uses this information to encourage candidates with similar profiles to apply for the job.[66] This selection means that a person that does not match this profile would be denied the opportunity to respond to this job offer due to them not being notified about it. A different example, mentioned by O'Neil, is the use of algorithms by HR departments to sift through job applicants to decide which person they will invite for an interview.[67] The algorithm performs a task, selecting qualified candidates, with a likely different outcome if this task was performed by a human. This requires not just regarding algorithms as dormant mathematical objects. It entails, as Mittelstadt et al. suggest, by seeing the algorithms as working within the computer and information systems they enable.[68]

In this paper I will follow the suggestion made by Mittelstadt et al., namely that a normative study of algorithms will have to take into account how these algorithms are implemented in computers systems and what tasks they enable to be performed.[69] Kitchin states that algorithms are "performative in nature and embedded in wider socio-technical assemblages"[70]. This means that algorithms cannot be seen separately from how they function in the real-world, as well as, how users respond to them and the consequences of the tasks they perform.[71] During normative research, algorithms should be studied as wider algorithmic systems, that include user-interaction, software programs, design decisions, databases, updates, and the tasks they perform. Assessing the employment of such algorithmic systems in the public sector has to expand this notion to also include the ability

---

[65] Kate Crawford and Jason Schultz, "Big data and due process: Toward a framework to redress predictive privacy harms." *BCL Rev.* 55 (2014): 99-101.
[66] Elizabeth MacBride. "How AI Aids Small Business Hiring: An Interview With ZipRecruiter's CEO." *Forbes*, October 31, 2017. https://www.forbes.com/sites/elizabethmacbride/2017/10/31/meet-the-jobs-startup-with-leverage-to-bring-google-and-facebook-to-the-table/2/#f6543d3235f7.
[67] O'Neil, *Weapons of Math Destruction*, 100-103
[68] Mittelstadt et al., " The ethics of algorithms: Mapping the debate" 2.
[69] Mittelstadt et al., " The ethics of algorithms: Mapping the debate" 2,3.
[70] Kitchin "Thinking critically about and researching algorithms" 16.
[71] Kitchin, "Thinking critically about and researching algorithms" 16, 19, 25.

for citizens to appeal against the actions of the algorithms or escape the grips of an algorithmic system, and whether or not the algorithmic systems produce the same consequences for different citizens in a consistent manner. The former is relevant on the grounds of the significant influence the actions of public institutions can have on the lives of individuals, and the latter to ascertain if the gains or impediments are divided equally and fairly. None of these aspects can be gleaned from merely studying the code. According to Kitchin, employing the same algorithm can lead to disparate outcomes in different situations. Kitchin states that: "algorithms perform in context – in collaboration with data, technologies, people, etc. under varying conditions – and therefore their effects unfold in contingent and relational ways, producing localised and situated outcomes”[72]. From the point raised by Kitchen it is clear that when looking at the employment of algorithms in the public sector a wide range of aspects needs to be taken into account. These include the choices regarding what is included or excluded from these datasets, how public servants interact with the algorithms, and how the algorithms are likely to change over time.[73]

Hill claims that the popular understanding of algorithms is slightly confused.[74] Although, it is relevant to make a distinction between the algorithm and the rest of the system, it is also necessary to take heed of how changing any individual component of the algorithmic system changes the consequences the implementation of the system has for public welfare. The consequences of the implementation of an algorithmic system might differ based on the choices or preferences of a public servant or the culture within a public institution. With the CAS-pilot this has been apparent, given that the same system functions in different ways across different police departments on the basis of a variety in how the end-users interact with the system.[75] It also makes sense to take a view of the entire system as a whole, because that is how citizens will experience it. For the case study of predictive policing it is not only relevant to merely look at the algorithm, the accuracy of the predictions or the content of the databases. It is also pertinent how the predictions are used in the real-world. This is especially true for predictive policing because, according to Perry, the predictions themselves are meaningless unless they are coupled with police action.[76] How the police officer or police official translates the prediction to action is therefore a substantive part of the system. This kind of interpretation would be closer to what Hill calls the popular understanding of algorithms. She claims that there is a difference between the mathematical objects called algorithms, and how the public understands and uses the concept of algorithms.[77] According to Mittelstadt et al., the popular usage focuses heavily on what is being done with the algorithms or what is enabled by them, rather than looking at what the

---

[72] Kitchin, “Thinking critically about and researching algorithms” 25.
[73] Kitchin, “Thinking critically about and researching algorithms” 16, 19, 25.
[74] Hill, “What an algorithm is” 36
[75] Bas Mali, Carla Bronkhosrt-Giesen, en Marielle Den Hengst.“Predictive policing: lessen voor de toekomst“ (2017). Accessed June 20, 2018. https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/93263.PDF. 32, 37, 41.
[76] Walt L Perry, *Predictive policing: The role of crime forecasting in law enforcement operations*. (Rand Corporation, 2013): 28.
[77] Hill, “What an algorithm is” 3.6

algorithm itself looks like.[78] They argue that it are exactly the tasks performed by algorithms, such as decision-making and classification, that lead to most of the ethical concerns.[79] Seeing algorithms in this way makes sense for assessing the employment of algorithms. Especially, considering that citizens will mostly experience an algorithm as constituted within the wider algorithmic system. Merely looking at the algorithm as the mathematical object would risk getting lost in the details.

[78] Mittelstadt et al., " The ethics of algorithms: Mapping the debate" 2.
[79] Mittelstadt et al., " The ethics of algorithms: Mapping the debate" 3.

## 2. Arguments for employment of algorithms

In this section I will give five arguments for the employment of algorithms in the public sector. The first argument is that data analysis by algorithms allows for proactive policy to be implemented. For instance, Achrekar et al. show that social media data can be used to preemptively provide medical care.[80] Secondly, the use of algorithms will lead to gains in efficiency, which will in turn lead to public institutions being able to provide more services to the public. Joseph and Johnson argue that automation will lead to gains through efficiency because there will be lesser reliance on public servant to perform task manually.[81] The third and fourth argument involve that employing algorithms allows public institutions to be more consistent and reliable in how public policy is enacted. This will be, in part, on grounds of human error and human bias becoming lesser of a liability in the implementation of public policy. Mayer-Schönberger and Cukier claim that is especially human cognitive bias to infer causal relations where there are none that opens for errors being made in policy-making.[82] Finally, I will address the argument that using algorithms leads to more objective public policy. For instance, Zarsky mentions that in law enforcement algorithms can be used to curtail the influence of human bias.[83]

### 2.1 New policy opportunities

The employment of algorithms in the public policy decision-making process will give rise to new opportunities for two distinct reasons. Firstly, the access to new kinds of information, that were previously difficult or neigh on impossible to measure, allow for informed decisions to be made on points where, due to a lack of knowledge, it would have been irresponsible to do so before. Secondly, it enables proactive policies to be implemented by public officials, in areas where they were at first limited to reactive measures. Social media data acts as a trove of data that was previously impossible to collect. Public officials might utilize such data to improve the effectiveness of their policies. For example, algorithms can be used to sort through vast amounts of social media data to aid public health officials to more effectively and efficiently respond to epidemics. Achrekar et al.[84], researchers specialized in computer sciences and medicine, describe how public health officials can use data derived from Twitter messages to quickly respond to influenza outbreaks. They found

[80] Achrekar et al., "Predicting flu trends using twitter data." 716, 717, 718.
[81] Joseph and Johnson, "Big data and transformational government." 46.
[82] Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data.* (New York: Houghton Mifflin Harcourt Publishing Company): 59-63.
[83] Tal Z. Zarsky, "Automated prediction: Perception, law, and policy." *Communications of the ACM* 55, no. 9 (2012): 35.
[84] Harshavardhan Achrekar, Avinash Gandhe , Ross Lazarus, Ssu-Hsin Yu and Benyuan Liu

that there is a notable correlation between Twitter messages that mention specific terms and actual flu cases.[85] Achrekar et al. mention that currently the process of tracking influenza outbreaks is to a large extent being done by manually processing the reports of healthcare professionals. This results in a delay of several weeks before the information becomes available to public healthcare officials.[86] Achrekar et al. suggest incorporating a system for analyzing data gathered from Twitter messages to aid public health officials in quickly responding to influenza outbreaks and provide preventative care.[87] This would involve using Twitter data to give priority to certain areas or risk groups when giving out influenza vaccinations. This would be far more efficient than having to react to when a person is already showing symptoms. Furthermore, it allows for protecting vulnerable groups from serious health risks.[88]

## 2.2 Efficiency

Yiu suggests that improving the use of data analytics in the public sector would allow the government to function more efficiently. He argues that increasing efficiency could lead to at least 16 billion pounds per year in savings for the British government.[89] The employment of algorithms in the public policy decision-making process will lead to more efficient outcomes for three reasons. Firstly, by employing algorithms parts of the decision-making process can be automated. This automation will result in the government agencies being able to be faster and cost-effective when providing services for citizens. Secondly, in some cases, algorithms will allow public policy-makers to allocate recourses preemptively. Thirdly, utilizing algorithms for data analysis can enable policy-makers to make more informed decisions on how and where to spend resources.

Rhoda Joseph and Norman Johnson argue that the employment of algorithms will lead to more efficient policies being implemented in the public sector. Joseph and Johnson, both specialized in Big Data, state that the use of algorithms by public institutions allows processes to be automated or redesigned in such a way that it would require considerably less time and recourses to provide services for, and complete transactions with, citizens.[90] A suggestion made by Joseph and Johnson is that the US Department of Veteran Affairs could complete the claims filed by US veterans quicker by moving away from a paper-based system and towards an automated system.[91] The handling of such claims would require far

[85] Achrekar et al., "Predicting flu trends using twitter data." 713
[86] Achrekar et al., "Predicting flu trends using twitter data." 713.
[87] Achrekar et al., "Predicting flu trends using twitter data" 716, 717, 718.
[88] Achrekar et al., "Predicting flu trends using twitter data" 716, 717, 718.
[89] Yiu, "The Big Data Opportunity" 6,7.
[90] Joseph and Johnson, "Big data and transformational government." 46.
[91] Joseph and Johnson, "Big data and transformational government." 45.

less effort and resources, by eliminating the need for civil servants to file each of them manually.[92]

Yiu mentions that by using advanced machine-learning algorithms public officials will be able to preemptively allocate resources to where they would be most effective.[93] As was seen in the example of using Twitter data to predict flu outbreaks, with some diseases it would be more efficient to preemptively vaccinate a person than it would be to treat the symptoms later on. Improving the implementation of data analytics will allow public servants to work more efficiently. For instance, Diakopoulos describes that the use of priorization algorithms has caused the housing inspection in New York to identify locations which would be most likely to have violations.[94] This results in housing inspectors being deployed more efficiently to such a degree that the amount of inspections that lead to an eviction notice has increased by 57 percentage points.[95]

## 2.3 Consistency

Decisions influenced by employing algorithms in the public sector will be more predictable and reliable for both public officials and the subjects of these decisions. Employing algorithms will aid in removing arbitrary or irrelevant aspects from the decision making process. For instance, the influence of the gut feeling of the loan officer or who the HR manager went to college with can be checked by introducing a reliance on algorithms that process relevant data sets. A scoring algorithm can be designed to disregard extraneous information in order to emphasize the aspects that would be most relevant for predicting future behavior or outcomes. They can even be designed to disregard information that might be relevant for making an accurate prediction, yet that is also highly prejudicial towards protected classes.

Algorithmic decision-makers remove other arbitrary aspects in the decision-making process. Blair et al.[96] show in a study on judicial sentencing that the facial of features of a defendant has an influence on the amount sentencing of a judge.[97] They found that judges award harsher sentences on African American defendants if those defendants have more African facial features. This is due to these facial features allowing for racial stereotypes to play a stronger role in the judicial decision-making process.[98] In a different study, Englich et al.[99]

[92] Joseph and Johnson, "Big data and transformational government." 46.
[93] Yiu, “The Big Data Opportunity”  15.
[94] Nicholas Diakopoulos, "Algorithmic-Accountability: the investigation of Black Boxes." *Tow Center for Digital Journalism* (2014): 4.
[95] Diakopoulos, “Algorithmic-Accountability: the investigation of Black Boxes." 4.
[96] Irene Blair, Charles Judd, and Kristine Chapleau
[97] Irene V. Blair, Charles M. Judd, and Kristine M. Chapleau, “The influence of Afrocentric facial features in criminal sentencing." *Psychological science* 15, no. 10, (2004): 674, 676.
[98] Blair et al.,  “The influence of Afrocentric facial features in criminal sentencing.678
[99] Birte Englich, Thomas Mussweiler, and Fritz Strack

have shown that the sentencing of judges is also strongly influenced by the sentencing demand of the prosecutor, even if it is clear that this was randomly generated.[100]

A risk assessment tool such as COMPAS or LSI-R can be used to alleviate concerns of irrelevant factors being decisive during sentencing procedures. Using algorithms to analyze input generated from tests done by a defendant can result in giving a risk-score.[101] This risk-score can then be used by the judge during the sentencing process to ensure a more fair sentencing record by reducing the influence of arbitrary factors and happenstance.

### 2.4 Reducing human error

The employment of algorithms in the public policy decision-making process will lead to fewer mistakes for two reasons. Firstly, automating tasks reduces the room for human error caused by cognitive biases or physiological elements. Secondly, Big Data analysis based on correlation eliminates the margin of error associated with sampling.

According to Viktor Mayer-Schönberger and Kenneth Cukier humans suffer from distinct cognitive biases that make us infer unfounded causal relations.[102] They state that: “such human intuiting of causality does not deepen our understanding of the world. In many instances, it's little more than a cognitive shortcut that gives us the illusion of insight but in reality, leaves us in the dark about the world around us”[103]. When humans act on these intuitions that produce misguided beliefs concerning causal links between phenomena, they will inevitably make errors.[104] For example, a referee in a soccer match might mistakenly decide to give a penalty to a team when he or she sees a striker fall in the penalty area after a challenge by a defender, even though there was not enough physical contact to make the striker fall. The human cognitive bias of the referee moves the referee to infer that enough contact between the defender and the striker must have taken place since something must have caused the striker to fall. The referee might be able to override this bias by relying on experience to tell that the striker took a voluntary dive, however, the point remains that human cognitive biases move us to make errors.

A second way humans are prone to make errors is the influence of physiological elements. For instance, Anderberg et al.[105] have shown that hunger makes a person more impulsive in

[100] Birte Englich, Thomas Mussweiler, and Fritz Strack, "Playing dice with criminal sentences: The influence of irrelevant anchors on experts’ judicial decision making." *Personality and Social Psychology Bulletin* 32, no. 2 (2006): 188, 194.
[101] Danielle, L. Kehl, Guo, Priscilla and Kessler, Samuel, “Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing”, *Berkman Klein Center for Internet & Society*, (2017): 11.
[102] Mayer-Schönberger and Cukier, *Big Data*, 59-63.
[103] Mayer-Schönberger and Cukier, *Big Data*, 63.
[104] Mayer-Schönberger and Cukier, *Big Data*, 63-66.
[105] Rozita Anderberg, Caroline Hansson, Maya Fenander, Jennifer Richard, Suzanne Dickson, Hans Nissbrandt, Filip Bergquist and Karolina Skibicka

his or her actions.[106] They claim impulsivity entails: "impaired decision making or action without foresight"[107]. This can be a source of errors in situations where impulsiveness is a liability or when it is combined with other physiological elements that prime us to make errors, such as sleep depravity.[108] For an individual, the impact will likely not go much further than increased expenses at the grocery store. Truly troubling cases come up when it affects decisions, made by public officials, that have far-reaching consequences on the lives of citizens. For instance, a public prosecutor might impulsively go for a higher sentence. As was shown in the previous section, such a decision can have a significant influence on the final jail time a defendant would have to serve.[109] The problem is aggravated because it introduces a degree of randomness to how public servants interact with citizens, that the citizens that are disadvantaged by this would understandably object to. Even more so in cases where the stakes are high, such as criminal sentencing.

An increased reliance on algorithms and automation will alleviate these problems. Firstly, due to algorithmic decision-makers being immune to such biases or physiological impediments. An algorithm would churn out the same result at 11 AM or 6 PM, and it would do so consistently given that all the relevant aspects remain the same. Furthermore, according to Gauri Naik and Sanika Bhide employing algorithms to aid in the diagnosis of diseases will reduce the mistakes made due to human error because algorithms are better at quickly analyzing data that is gathered with the use of microscopes or specialized medical scanners.[110] Secondly, according to Mayer-Schönberger and Cukier, implementing analysis based on Big Data practices will lead to better decisions being made.[111] They describe that extrapolating from sample sizes necessarily leads to errors in the final result. With the rise of Big Data, larger data sets can be analyzed so that there will be no need for sampling. This means that the margin of error associated with sample sizes can be eliminated.[112] Although Mayer-Schönberger and Cukier state that these data sets are not necessarily large enough to be beyond what humans would be able to analyze, however, in many cases they will be.[113] Mayer-Schönberger and Cukier mention the classic example of how chess algorithms have fully solved certain game states in chess. According to Mayer-Schönberger and Cukier, this is to such a degree that humans would never be able to beat these computer systems.[114] Poignantly, in the case of public policy-making, the employment of algorithms to sift through large data sets will lead to gaining valuable insights that would not be attainable using

---

[106] Anderberg, Rozita H., Caroline Hansson, Maya Fenander, Jennifer E. Richard, Suzanne L. Dickson, Hans Nissbrandt, Filip Bergquist, and Karolina P. Skibicka. "The stomach-derived hormone ghrelin increases impulsive behavior." *Neuropsychopharmacology* 41, no. 5 (2016): 1199-1120.
[107] Anderberg et al., "The stomach-derived hormone ghrelin increases impulsive behavior." 1199.
[108] Junu Pilcher,, and Allen I. Huffcutt. "Effects of sleep deprivation on performance: a meta-analysis." *Sleep* 19, no. 4 (1996): 319.
[109] Englich et al, "Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making."196.
[110] Naik and Bhide, "Will the future of knowledge work automation transform personalized medicine?" 51
[111] Mayer-Schönberger and Cukier, *Big Data*, 59.
[112] Mayer-Schönberger and Cukier, *Big Data*, 26- 33.
[113] Mayer-Schönberger and Cukier, *Big Data*, 31, 37, 54, 103, 177.
[114] Mayer-Schönberger and Cukier, *Big Data*, 31.

traditional methods. Policy-makers will, in turn, be able to make better decisions using these insights, as well as, balancing out their personal biases or unchecked preconceived notions.

## 2.5 Objectivity

The employment of algorithms can aid in reducing the influence of human bias in the decision-making process. This can be achieved by investing in automation, reducing the total amount of decisions that need to made by humans or by providing a tool to aid human-decision makers. This tool might be a score or model based on objective criteria generated with the use of algorithms.

Gillespie describes that algorithms are widely regarded as: "stabilizers of trust, practical and symbolic assurances that their evaluations are fair and accurate, free from subjectivity, error, or attempted influence".[115] For instance, Naik and Bhide state that Optra Health has created software that can lead to an "unbiased and speedy diagnosis"[116] of diseases by employing algorithms to process data gathered from microscopes or scanning equipment.[117] Tal Zarsky, a legal scholar, makes a moral argument for the employment of algorithms in law enforcement. He argues that hidden biases cause the decisions made by law enforcement officials to be systematically detrimental to minorities or vulnerable groups that have long been subjected to the effects of strong prejudices in a society.[118] These hidden biases can subconsciously influence the decision made by a law enforcement official, in the same way, as it did in the loan application example. Zarsky claims that: "limiting the role of human discretion and intuition and relying upon computer-driven decisions this process protects minorities and other weaker groups"[119]. This can be achieved by either automating the decision-making and risk-assessment process or by employing algorithms to sort through more data, giving public officials access to more objective information and reducing the need to rely on their intuition.[120]

Gernot Rieder and Judith Simon, specialized in computer ethics and public policy, claim that this connection between algorithms and objectivity taps into a historic tradition of epistemic virtues.[121] They state that the rise of Big Data builds on older epistemic virtues by promising to eliminate human interference and in effect human bias, from the decision-making process. This can be achieved with increased automation of the decision-making process, reducing the role played by humans in the decision-making process, or by relying on

[115] Gillespie, "The relevance of algorithms" 13.
[116] Naik and Bhide, "Will the future of knowledge work automation transform personalized medicine?" 51.
[117] Naik and Bhide, "Will the future of knowledge work automation transform personalized medicine?" 51, 53.
[118] Zarsky, "Automated prediction: Perception, law, and policy," 35.
[119] Zarsky, "Automated prediction: Perception, law, and policy," 35.
[120] Zarsky, "Automated prediction: Perception, law, and policy," 33, 34, 35.
[121] Rieder Gernot, and Judith Simon. "Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data." *Big Data & Society* 3, no. 1 (2016): 2.

employing predictive techniques to: “support decision making and optimize resource allocation across many government sectors. Applying a mechanical mindset to the colonization of the future.”[122]. This mechanical mindset would also be released on data that was previously difficult to measure, such as social media data, making objective knowledge available about new domains.[123]

Algorithmic scoring systems can also be used as a tool to aid human decision-makers to reduce the influence of human bias. During the development, process algorithms can be set up in such a manner that they disregard any information that would either be irrelevant for predicting future behavior or preferably would not be taken in consideration on legal or normative grounds. With the FICO-score example, O’Neil describes that factors such as race or gender were often leading in whether or not a person had access to a line of credit while they were not necessarily relevant for predicting future behavior. The FICO-score algorithm disregards any such biographical information in favor of more salient and objectively measurable criteria such as monthly earnings, expenditures, and past payment behavior.[124] A scoring algorithm can be used as a tool to aid a human decision-maker to come to more objective decisions by checking the loan officer’s biases. A scoring algorithm can also be used to fully automate the loan application process, removing the need for direct human influence altogether. Another example of the employment of scoring algorithms is the use of risk scoring algorithms used by judges during sentencing procedures, such as COMPAS and LSI-R.[125] By designing algorithms to only take into account information that is directly relevant for predicting future behavior or outcomes subjective criteria and prejudices associated with human judgment can be warded off from the decision-making process.

[122] Rieder & Simon, “Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data,” 4.
[123] Rieder& Simon, “Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data,” 4.
[124] O'Neil, *Weapons of Math Destruction*, 142, 143.
[125] Kehl et al., “Algorithms in the Criminal Justice System” 11.

## 3. Arguments against the employment of algorithms

In this section, I will give five counterarguments to the arguments made in the previous section. Each of these refutations will address a specific argument made in favor of the reliance on algorithms in the public sector. The first counterargument is intended to refute the notion of algorithms being entirely neutral or objective entities. For instance, Kraemer et al., show that taking a stand with regards to ethical problems occasionally will be a necessary step during the development process of an algorithm.[126] The second counterargument shows that there are limitations to the gains that can be achieved through maximizing for efficiency. From the arguments made by Haskins it can even be derived that a small margin of inefficiency is necessary for public institutions to function.[127] The third argument challenges the notion that consistency is to be regarded as entirely beneficial. From the example of the dominance of hiring algorithms given by O'Neil it can be seen that at a certain point the logic of an algorithmic system can become ubiquitous, making its rules inescapable. The result will be that the collateral damage caused by the use of an algorithm will be unduly concentrated towards specific groups.[128] Fourthly, although the employment of algorithms will markedly reduce human-error, they introduce new kinds of mistakes. According to Diakopoulos, classification algorithms inevitably churn out either false positives or false negatives.[129] Combined with the notion of algorithms being objective this can cause citizens to be trapped in the vise of an algorithmic system. Finally, I will argue that given these problems with the employment of algorithms it is necessary to maintain a way for public servants to break the loop of an algorithmic system. This can be done by, as argued for by Brenninkmeijer, opening up a pathway for citizens to appeal to a public servant with sufficient discretionary power to bend the rules or by building in a role for public servants to supervise and correct the actions performed by algorithms.[130] Both these options will necessarily limit how far an algorithmic system can be implemented.

### 3.1 Bias in algorithms

The notion that algorithms are free from bias can be crucial for the legitimacy of an organization. According to Gillespie, this is especially true for algorithms that decide what

---

[126] Kraemer, et al., "Is there an ethics of algorithms?" 255, 258.
[127] Haskins, "The drawbacks of efficiency"
[128] O'Neil, *Weapons of Math Destruction*, 100
[129] Diakopoulos, "Algorithmic-Accountability: the investigation of Black Boxes." 6
[130] Brenninmeijer, Alex. "Meer dan een burger-servicenummer." Accessed June 20, 2018. https://www.nationaleombudsman.nl/uploads/tijdschrift_voor_conflicthantering_hoeveel_legitimatie_schept_lokale_democratie.pdf , 10.

kind of information is shown to whom.[131] He states that: “Above all else, the providers of information algorithms must assert that their algorithm is impartial. The performance of *algorithmic objectivity* has become fundamental to the maintenance of these tools as legitimate brokers of relevant knowledge. No provider has been more adamant about the neutrality of its algorithm than Google, which regularly responds to requests to alter their search results with the assertion that the algorithm must not be tampered with”[132]. However, Gillespie argues, information providers necessarily have to make countless evaluations with regards to how their algorithms function in the real world. In a different article Gillespie shows that information providers use this selection of information as a selling point to promote ad-space.[133] Algorithms cannot be seen separately from the tasks they perform, and the consequences derived from these actions. Even more so, if an organization is called upon to interfere with the workings of an algorithm. Gillespie shows that Google has altered their algorithm due to political concerns. In 2009 Google felt forced to intercede upon the workings of their algorithm after a white supremacist group created an offensive image Michelle Obama and made it a top-ranking image search result.[134] Strikingly, the subjective influence of Google also affects their algorithm by virtue of when Google chooses not to interfere with their algorithm. Google drew the line with the image of Michelle Obama, because the image was immediately visible as a thumbnail in the search result.[135] The choice where to draw the line, when interferences with the algorithm becomes necessary, cannot be solely based upon objective criteria. This is a subjective choice on the part of the user of the algorithm. Given that these choices might be influenced by human bias, the way that algorithms function cannot be considered entirely objective.

The notion that algorithms are free from bias is being widely challenged. For instance, Crawford and Schultz claim that: “One of the primary myths about Big Data is that it produces outputs that are somehow free from bias and closer to objective truth than other forms of knowledge”[136]. Friedman and Nissenbaum even differentiate between three categories of bias in computer systems. They describe that computer systems are biased when they: “*systematically* and *unfairly discriminate* against certain individuals or groups of individuals in favor of others. A system discriminates unfairly if it denies an opportunity or a good or if it assigns an undesirable outcome to an individual or group of individuals on grounds that are unreasonable or inappropriate.”[137] The first category of bias is pre-existing bias. Friedman and Nissenbaum say that pre-existing bias “is when computer systems embody biases that exist independently, and usually prior to the creation of the system, then we say that the system embodies pre-existing bias”[138]. This kind of bias can either

131 Gillespie, “The relevance of algorithms“ 12.
132 Gillespie, “The relevance of algorithms“ 14.
133 Gillespie, “The relevance of algorithms“ 14.
134 Gillespie, “The relevance of algorithms“ 15.
135 Gillespie, “The relevance of algorithms“ 15.
136 Crawford & Schultz,“ Big data and due process” 127.
137 Batya Friedman and Helen Nissenbaum, "Bias in computer systems." *ACM Transactions on Information Systems (TOIS)*,14, no. 3 (1996): 332.
138 Friedman and Nissenbaum, “Bias in Computer Systems” 333.

originate from the bias of an individual designer or it can be inherited from the reining biases in a society or institution.[139] The second category of bias in computer systems is technical bias. According to Friedman and Nissenbaum, technical bias comes from technical limitations or constraints.[140] Such constraints might be due to the process of translating human constructs to input for algorithms, imperfections in the randomization process, algorithms that fail “to treat all groups fairly under all significant conditions”[141], and limitations caused by the technology available itself. The latter being limitations to how much data would be economically feasible to store or process given the hardware that is currently on the market.[142] Technical constraints also force design choices to be made, these choices are markedly subjective. The third category mentioned by Friedman and Nissenbaum is emergent bias. This category specifically addresses the kind of bias that emerges through interaction with users.[143] Friedman and Nissenbaum argue this might happen when a computer system has become outdated due to the “emergence of new knowledge in society that is not or cannot be incorporated in the system design”[144] or when a computer system is used by users with different knowledge and expertise than was calculated for during the design process.[145]

Given these three categories of bias, it is also relevant to further identify how and when bias is introduced into an algorithm. There are five distinct ways of how or when bias can creep into an algorithm. The first is the choices made by the designer during the development process. Vedder and Naudts state that: “One of these elements can be the bias of an algorithm’s designer. Algorithms are human constructs and therefore the algorithmic process will almost necessarily assume, perhaps unwillingly, certain values”[146]. They further argue that the consequences of these choices made by designers are difficult to predict and ascertain due to the complexity of the context algorithms operate in, especially in a real-world context or when they start affecting human lives.[147] Examples of subjective criteria that play a role in the development process might be preferences of the designer or assumptions about the desired end-users.

Mittelstadt et al. argue that algorithms necessarily inherit the biases of the humans that are involved in the development process. They claim that: “Algorithms inevitably make biased decisions. An algorithms design and functionality reflects the values of its designers. Functionality also reflects the values of designers and intended users, if only to the extent that a particular design is preferred as the best or most efficient option.”[148]. It is the clients,

139 Friedman and Nissenbaum, “Bias in Computer Systems” 333, 334.
140 Friedman and Nissenbaum, “Bias in Computer Systems” 334.
141 Friedman and Nissenbaum, “Bias in Computer Systems” 334.
142 Friedman and Nissenbaum, “Bias in Computer Systems” 334.
143 Friedman and Nissenbaum, “Bias in Computer Systems” 335.
144 Friedman and Nissenbaum, “Bias in Computer Systems” 335.
145 Friedman and Nissenbaum, “Bias in Computer Systems” 335.
146 Anton Vedder and Laurens Naudts. "Accountability for the use of algorithms in a big data environment." *International Review of Law, Computers & Technology* 31, no. 2 (2017): 206-224. 208
147 Vedder and Naudts, “Accountability for the use of algorithms in a big data environment” 208, 209, 210.
148 Mittelstadt et al., “The ethics of algorithms: Mapping the debate” 7.

in the public sector the policy-makers, that determine what task the algorithm is to perform, and what would be an acceptable way of achieving this. Mittelstadt et al claim it is then up to the designer to determine how to create an algorithm that fits the specifications given by the client or end-user. The designer will have to make choices during the development process. The design choices matter for how the algorithm, and the corresponding computer system, functions. Mittelstadt et al. show that these choices are based on subjective criteria. This makes the resulting algorithm a matter of contingency. Therefore, a difference in design choices causes a difference in outcomes. [149]

With program synthesis or automatic programming, it is possible that algorithms are creating new algorithms.[150] This would circumvent the point, made by Mittelstadt et al., of human expertise or human design choices being a determining factor in the resulting algorithm as well as the outcomes generated from the implementation of the algorithm.[151] However, the algorithm still is influenced by human choices made during the design process. The client has to formulate the instructions about what needs to be achieved, and what the acceptable parameters are for achieving these goals. [152] Furthermore, given that, as stated by Kitchin, algorithms are ontogenetic in nature they would be influenced by humans through updates, even if they were free from bias at inception.[153]

Kraemer et al. show how the design process of algorithms might involve taking a stand on ethical and philosophical questions. This is due to the possibility that designers with different views on ethical questions would make different choices during the design process, based on those convictions. They claim that an algorithm: “comprises an essential value-judgment if and only if, everything else being equal, software designers who accept different value-judgments would have a rational reason to design the algorithm differently (or choose different algorithms for solving the same problem)”[154]. Kraemer et al. illustrate this point with an example of the use of classification algorithms in medical practice.[155] They argue that with medical diagnosis software there necessarily will be a tradeoff between achieving false positives or false negatives. Setting a lower threshold leads to more false positives, however, it also reduces the risk of an illness reaming undetected. Contrariwise, setting a higher bar for the algorithm to classify a persons as having a disease would lower the chance a person falsely gets flagged as having a disease while it might also cause diseases to be discovered at a later time or missed entirely.[156] False positives cause unnecessary stress for patients and increased healthcare costs while false negatives lead to serious health risks due to more treatment options being available when a disease is discovered at an earlier stage.

---

[149] Mittelstadt et al., “The ethics of algorithms: Mapping the debate” 7.
[150] Ali, Tariq. “Will programming be automated? (A Slack Chat and Commentary)” Accessed June 21, 2018. https://dev.to/tra/will-programming-be-automated-a-slack-chat.
[151] Mittelstadt et al., “The ethics of algorithms: Mapping the debate” 7.
[152] James Bomholt, “Program Synthesis Explained.” Accessed June 20, 2018. https://homes.cs.washington.edu/~bornholt/post/synthesis-explained.html
[153] Kitchin “Thinking critically about and researching algorithms” 18, 21.
[154] Kraemer, et al., “Is there an ethics of algorithms?" 253.
[155] Kraemer, et al., “Is there an ethics of algorithms?" 253,254, 256, 258.
[156] Kraemer, et al., “Is there an ethics of algorithms?" 255.

According to Kraemer et al., there is no objective way to decide how to strike a balance between skewing the system towards producing either false positives or false negatives. Therefore, they argue, the choice "for a certain threshold in an algorithm is a decision that is a judgment about which there is, or at least could be, a controversy between advocates of the major theories of normative ethics. This supports our claim that algorithms manifest or reflect certain ethical judgments. This point can be illustrated as a choice between deontological and consequentialist or utilitarian theories of normative ethics."[157]. They claim that from a Kantian, or deontological, perspective the emphasis will likely be placed on the "physical and mental integrity of the individual patient"[158], thereby, designing the algorithm towards producing more false positives while a consequentialist approach would emphasize the integrity of the data that is being collected to ensure better long-term results.[159] This means that designers might not only make different choices based on adherence to different normative theories, they might further differ based on diverging interpretations of such theories. From the example given by Kraemer et al. it is clear that normative values can play a determining role in the design process. In some cases this will even be necessarily so due to the unavailability of objective criteria. The resulting algorithm can, therefore, not be considered free from bias nor wholly objective.[160]

A second source for bias to creep into algorithms is from the values and prejudices of the client. The influence of the client start from the decisions made during policy design and will be felt throughout the enactment of the policy. The client provides instructions on what needs to be achieved. This means that the client formulates what the desired outcome would be that the computer system is meant to aid in setting about. O'Neil argues that how a model functions is strongly predicated by the measurement of success the user formulates beforehand.[161] How success is defined is dependent on the values that are being highlighted. For instance, O'Neil states that there is a strong difference between how she or her children would define a successful meal. The values that are leading in how she defines success are the nutritional value of the meal and time needed to prepare the meal while her children would want to optimize the amount of Nutella.[162] The same argument can be made about how algorithms and computer systems function. It depends on those that are the responsible for making the policy to formulate what is considered to be a successful outcome. This involves the making of choices of which values are being highlighted and how tradeoffs will be resolved. For instance, when there are tradeoffs between gains through efficiency and the privacy of citizens the set of instructions embedded in the algorithm will reflect the values of the client and which of these conflicting values the client prioritizes. These values will not only be leading during the design process. Systems that succeed in

[157] Kraemer, et al., "Is there an ethics of algorithms?" 258.
[158] Kraemer, et al., "Is there an ethics of algorithms?" 258.
[159] Kraemer, et al., "Is there an ethics of algorithms?" 258.
[160] Kraemer, et al., "Is there an ethics of algorithms?" 253, 254-258
[161] O'Neil, *Weapons of Math Destruction*, 27, 51, 129, 180.
[162] O'Neil, Cathy. "Interview with Cathy O'Neil, author, Weapons of Math Destruction" , *Beyond the bookcast* , November 14, 2016. Accessed June 19, 2018. http://beyondthebookcast.com/wp-content/uploads/2016/11/ONeilTranscript.pdf.

optimizing the values of a client are likely to be implemented while those that do not will be discontinued. How success is defined by the client will be a determining factor for the consequences the implementation of an algorithm will have. This is especially relevant for computer systems implemented in the public sector, because these can have far-reaching consequences on the lives of individual citizens.

The third way bias can be introduced into how algorithms function corresponds with the category of technical bias.[163] For example, the amount of information that can be shown on a screen is a hardware limitation the designer of an algorithm needs to take into account. According to Friedman and Nissenbaum, the limitation of the number of search results that can be shown at the same time on a screen makes the results on the first few screens significantly more visible to users. This causes those donor candidates shown on the first screen to have a higher chance to receive a donor organ.[164] Therefore, Friedman and Nissenbaum argue, ranking algorithms can be biased if they systematically favor certain groups over others, by placing them earlier in the search results.[165]

A fourth source of bias for algorithms are the data sets. Bias can creep into algorithms through the data sets that are fed into the algorithm. This can either be from the bias that is present in the data set or choices made regarding which data to include or exclude. With the example of credit-scores, Pasquale argues that data sets filled with bias and prejudice data get laundered into seemingly objective data sets.[166] He states that as “Subtle but persistent racism, arising out of implicit bias or other factors, may have influenced *past* terms of credit”[167], and given that these prejudices caused specific groups to pay higher interest rates, the data set would show that applicants from those groups have a higher chance to default on their loans. This makes sense because it is more difficult to pay back a loan when the interest rates swallow a larger portion of a person’s monthly income. According to Pasquale, this becomes problematic when data sets that contain historic bias are fed into algorithms to decide the interest rates of future applicants. The algorithm would systematically give applicants from certain groups higher risk-scores. Applicants from these groups would, subsequently, be more likely to default on their loan, due to these higher interest rates. This leads to fresh, and seemingly neutral, data predicting that an applicant from that group should be awarded a higher interest rate, given that they match the profile of other applicants that reneged on paying back the loan in time.[168]

O’Neil describes a similar problem with risk-scoring algorithms used during sentencing, such as COMPAS or LSI-R. A person might have a high risk-score based on that person matching the profile of other defendants that recidivated. When a judge makes the decision to give a

---

[163] Friedman and Nissenbaum, “Bias in Computer Systems” 334, 335.
[164] Friedman and Nissenbaum, “Bias in Computer Systems” 334.
[165] Friedman and Nissenbaum, “Bias in Computer Systems” 334.
[166] Pasquale, *The Black Box Society*, 41.
[167] Pasquale, *The Black Box Society*, 41.
[168] Pasquale, *The Black Box Society*, 41.

longer sentence to that person based on this risk-score a “pernicious feedback loop”[169] might occur. According to O’Neil, the long jail time and criminal record will severely impede the chances that this person will be able to get a job after serving his or her sentence. This lack of employment makes it more likely that the person will commit another crime, thereby, reinforcing the prediction made by the risk-scoring model and the next defendant with a matching background will have an ever higher risk-score.[170]

Solon Barocas and Andrew Selbst, legal scholars in an article about algorithms and the rights of protected classes, discuss how algorithms used during the hiring process might enhance and systematize the biases of employers.[171] They claim that if an employer has shown a bias towards a certain type of candidate, the matching algorithms will learn from this historic hiring data to recommend candidates that match this profile. More importantly, the algorithm will also learn not to recommend candidates that match the profile of previous candidates the employer chose not to hire. They argue that: “if LinkedIn's algorithm observes that employers disfavor certain candidates who are members of a protected class, Talent Match may decrease the rate at which it recommends these candidates to employers. The recommendation engine would learn to cater to the prejudicial preferences of employers.”[172]. Thereby, the employing of matching algorithms is leading employers to unwittingly make biased hiring decisions based on the prejudices of their predecessors.[173] The algorithm used by ZipRecruiter enhances these problems by learning from the hiring decisions made by an employer which applicants will be shown job openings. The algorithm will decide not to show a specific job opening to a person if that person does not match the profile of applicants that were previously hired by that employer.[174] This will cause groups to be denied the opportunity to apply for the job, because of biases that might have played a role in past hiring practices.

This will not only be detrimental to the candidates who are being overlooked due to the influence of pre-existing biases, it might also hamper the performance of public sector institutions. For specific public sector organizations a lack of diversity in the workforce impedes on the quality and efficiency of the work being done. With policing there are distinct advantages in being able to reflect the diversity of the community that the police force is serving. For instance, the hiring of openly gay police officers has proven to be a boon in.[175] This does not mean that a policy of promoting hiring diversity is a catch-all solution, however, clearly biased hiring practices will harm the ability of a police force to effectively administer its duties.

---

[169] O'Neil, *Weapons of Math Destruction*, 20, 35.
[170] O'Neil, *Weapons of Math Destruction*, 32-36.
[171] Solon Barocas and Andrew D. Selbst, “Big data's disparate impact.” *California Law Review*, 104, (2016): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899: 683
[172] Barocas and Selbst, “Big data's disparate impact.” 683.
[173] Barocas and Selbst, “Big data's disparate impact.” 683.
[174] ElizabethMacBride, “How AI Aids Small Business Hiring: An Interview With ZipRecruiter's CEO.”
[175] Donatella Lorch. “Openly Gay in Blue: Officers Tread Warily.”*The New York* Times, July 13, 1992. https://www.nytimes.com/1992/07/13/nyregion/openly-gay-in-blue-officers-tread-warily.html.

The fifth way of how and when bias can creep into algorithms is through user interaction. Engin Bozdag, in an article about bias in algorithmic filtering and personalization, claims that this can happen after a computer system has been implemented.[176] This is what Friedman and Nissenbaum call emergent bias.[177] They argue that emergent bias is a problem that arises from how users interact with computer systems. The way that users interact with a computer system is mediated by the user interface (UI). The design of a UI involves making choices between competing values. For instance, a complex UI will give more options to users while it can also make it exceedingly difficult for inexperienced users to learn how to navigate the UI. Furthermore, a complex UI will increase the likelihood of users making mistakes or engage in interactions that they not want to do or fully understand the consequences of.[178] According to Nissenbaum and Friedman, problems of emergent bias arise when the intended users during the design phase of a computer system are notably different from the actual or future users. This means that far beyond the development process an algorithm remains vulnerable to picking up biases based on choices made in the past. [179]

When the intended audience during the development of the computer system has been users with a high digital literacy or mastery of the Dutch language the designer will make different choices than when the client emphasis that the system should be easily accessible to a wide range of users. For the public sector this will revolve around whether or not the public servants are capable of sufficiently navigating the UI. A gap between the actual capabilities of the civil servant and the digital literacy required to successfully interact with the computer system will cause public servants to be unable to do their work properly. Such a gap might prevent a civil servant to get access to all the relevant information due to an inability to understand the way the information is presented by the UI. For citizens, the UI can become a unsurpassable barrier to communicate with governmental agencies and get access to services. Groups with lower digital literacy will find it more difficult to get access to governmental services because they are unable to successfully interact with the computer system. This will entail that groups with lower digital literacy will be systematically unable to get the services they are entitled to due to earlier made design choices. This can aggravate already existing inequalities. Moreover, language is an important part of the UI. If the UI was designed with a certain level of language skill in mind, and specific groups are unable to fully understand the text, these users will be prone to make mistakes in their interactions with public sector institutions. This opens these groups up to be penalized. Merely adding an English option will not be a solution if the most vulnerable groups do not understand formal English or have difficulty reading in general. Alex Brenninkmeijer, also states that the

[176] Bozdag, Engin."Bias in algorithmic filtering and personalization." *Ethics and information technology*, 15 no. 3, (2013): 210.
[177] Friedman and Nissenbaum, "Bias in Computer Systems" 335, 336.
[178] The Interaction Design Foundation. "User Interface (UI) Design", *The Interaction Design Foundation*. Accessed June 19, 2018.https://www.interaction-design.org/literature/topics/ui-design
[179] Friedman and Nissenbaum, "Bias in Computer Systems" 335.

overestimation of the ability citizens have to understand complex digital systems is a major point of concern, and significant source of frustration for Dutch citizens.[180]

The emergent biases Friedman and Nissenbaum discuss are mostly unintentional side products due to a shift of user demographics or a general lack of foresight in the development process.[181] However, with UI design there are also ethical concerns when designers intentionally exploit the gap between the capabilities of the user and the required digital literacy to successfully navigate the UI. This can be done by increasing the complexity of the UI or the use of dark patterns.[182] Dark patterns are stratagems in UI design meant to make a user interact with a computer system in an unintended manner.[183] In the public sector, dark patterns, combined with nudging, have proven to be a toxic combination that cause severe problems for vulnerable groups, such as migrants.[184] The point is that algorithms cannot be seen separately from the way they function in the real world. Especially when algorithmic predictions need to be understood by the public servants that base their decisions on them. Furthermore, automated decision-making algorithms will penalize groups that are unable to successfully interact with them. Complex UI's will be a barrier to both of these issues. This means that, as argued by Friedman and Nissenbaum, bias can creep into algorithms through user interaction.[185]

Far from being entirely free from bias, algorithms are influenced by subjective factors in several ways. Firstly, bias can originate from the design choices made during the development process. Kraemer et al. show that in some cases the designer of an algorithm is forced to take a stand with regards to ethical questions.[186] Poignantly, how the algorithm functions might vary based on the values of the designer. In the example of medical diagnosis software the algorithm will favor producing either false positives or false negatives.[187] This will have significant consequences for both the patient and the workload of healthcare providers. Secondly, the instructions given by the client are also determinant for how an algorithm functions. O'Neil argues that a client introduces a subjective influence into a model by formulating a measurement for success.[188] For algorithms, the client will formulate what the computer system is to bring about, and what criteria are to be measured to determine this goal has been realized. Some of these choices will depend on the values of the client, especially in situations where a choice needs to be made between competing values. How such a conflict is resolved cannot be regarded as solely being based on objective criteria. Thirdly, technical bias arises when the way algorithms function combines with

[180] Brenninkmeijer, "Meer dan een burger-servicenummer"
[181] Friedman and Nissenbaum, "Bias in Computer Systems" 335.
[182] UIE. "Dark Patterns and the Ethics of Design" *Adventures in UX design*, 19, November 29, 2017. https://medium.com/adventures-in-ux-design/dark-patterns-and-the-ethics-of-design-31853436176b
[183] Dark Patterns "What are Dark Patterns?" Accessed June 18 2018. https://darkpatterns.org/
[184] Barret, Nick. "Hostile environment: The dark side of nudge theory." Accessed June 21, 2018. http://www.politics.co.uk/comment-analysis/2018/05/01/hostile-environment-the-dark-side-of-nudge-theory
[185] Friedman and Nissenbaum, "Bias in Computer Systems" 335.
[186] Kraemer, et al., "Is there an ethics of algorithms?" 258.
[187] Kraemer, et al., "Is there an ethics of algorithms?" 255, 258.
[188] O'Neil, *Weapons of Math Destruction*, 25.

hardware limitations in such a way that it systematically disadvantages specific groups. A fourth source for bias in algorithms are the data sets that are fed into the algorithm. According to Pasquale, data sets might reflect the historic prejudices of a society or individual decision-makers. When an algorithm is fed these data sets it will acquire the contained prejudices and reinforce them.[189] Finally, bias can be introduced after the implementation of an algorithm. According to Friedman and Nissenbaum, emergent bias comes about when a computer system is built for a different user base than the actual or future users of the system.[190] This is relevant for public policy making if contact between citizens and governmental agencies is mediated through computer systems. When certain groups of citizens are faced with overly complex UI's they will be unable to get access to governmental services or face to be unfairly penalized by automatic decision-making algorithms.

### 3.2 Inflexibility and Exclusion

As was seen in section 2.2, the employment of algorithms allows for more efficient use of resources in the implementation of public sector policy. For example, the automation of interaction between citizens and governmental agencies will reduce the amount of time and resources required to provide services to citizens.

Casey Haskins, director of the Department of Military Instructions at U.S. Military Academy at West Point, warns that too much efficiency is detrimental to the overall functioning of governmental organizations.[191] He does not deny that efficiency is an important good that public sector institutions should strive for. However, Haskins argues, maintaining some degree of inefficiency provides two crucial benefits. Firstly, an organization is more resilient when it has some reserves that it can tap into during a crisis.[192] Retaining reserves makes an organization cumbersome, and likely to be anathema to a manager that wants the organization to be as lean and agile as possible, however, these reserves are vital when there is an unforeseen spike in demand. Haskins illustrates his point with stating that a hospital would do well to have some overcapacity in available hospital beds, to avoid being entirely overrun when a pandemic breaks out.[193] The second advantage of inefficiency for institutions mentioned by Haskins is that inefficiency makes an organization more adaptable.[194] When an organization facilitates some degree of experimentation the organization will be able to gain valuable insights. Haskins argues these insights are valuable assets for an organization when it needs to respond to changing circumstances. Experimentation necessarily involves giving employees leeway in how they perform their

---

189 Pasquale, *The Black Box Society*, 41.
190 Friedman and Nissenbaum, "Bias in Computer Systems" 335
191 Haskins, "The drawbacks of efficiency"
192 Haskins, "The drawbacks of efficiency"
193 Haskins, "The drawbacks of efficiency"
194 Haskins, "The drawbacks of efficiency"

tasks. This leeway comes at the cost of inefficiency by making it more difficult to eliminate human made errors. According to Haskins, the necessity for government organizations to be able to keep up with changing circumstances, combined with the importance of having access to reserves in times of need, makes preserving a small amount of inefficiency crucial for an organization to function.[195]

Efficiency through the employment of algorithms can be detrimental to citizens on the bases of the value of public servants retaining some degree of flexibility in how they interact with citizens. This is most apparent in situations where exceptions need to be made in order to avoid absurd or unfair outcomes. If the U.S. VA department would follow through on the suggestion made by Joseph and Johnson, to fully automate the filing process, it would lead to an overall gain in efficiency and smaller waiting period for U.S. veterans to receive their benefits, however, by completely removing human involvement from the process there is no clear possibility for human intervention in cases where the system goes awry. Resolving problems with corner cases will often require bending the rules. In the public sector, being able to resolve these problems is of great importance on the grounds of the vast consequences these problems will have for citizens. It is not a matter of not being able to get the preferred model of phone or favorite brand of beverage, it is about receiving social benefits or jail time. Even with credit ratings there are possibilities to go to different service providers, such as Upstart[196]. With governmental services there is very little leeway or choice on the part of the citizens. Therefore, there has to be more flexibility built in on the side of the public sector institution. As was seen with the arguments made by Haskins, maximizing efficiency at some point invariably will come at the expense of flexibility. Given the vast consequences of the actions made by public institutions there is a limit to how much can be automated.

Besides problems associated with inflexibility, the employment of algorithms in the decision-making process of public institutions might also lead to unreliable predictions on the grounds of Big Data exclusion. According to Jonas Lerman, specific groups are likely to be underrepresented in data sets. He states that it are especially members of disadvantaged groups that leave smaller data trails.[197] Indeed, it makes sense that the data collected on the basis of Amazon prime memberships or other online subscriptions would favor the more affluent. Likewise, such overrepresentation can be found in data derived from online shopping behavior or social media. Lerman raises ethical concerns over the use of information based on algorithms data mining data sets wherein specific groups are structurally underrepresented.[198] He states: “These technologies may create a new kind of voicelessness, where certain groups' preferences and behaviors receive little or no consideration when powerful actors decide how to distribute goods and services and how to

[195] Haskins, “The drawbacks of efficiency”
[196] Fogg, Kathleen. “What Predicts Loan Repayment at Auto Capital?”, Honors Project, Spring, (2016): 1-28 13
[197] Jonas Lerman, “Big data and its exclusions.” *Stanford Law Review*, Online, 66, (2013):58, 59.
[198] Lerman, “Big data and its exclusions.” 58,59, 60.

reform public and private institutions.”.[199] Lerman argues that this is especially relevant when such data sets are used for the allocation of resources by public institutions.[200] In section 4.2, I will argue, that this goes against the egalitarian stride, that follows from welfare utilitarianism, described by Goodin.[201] Given that Lerman is correct, that it are indeed the underprivileged groups that will be overlooked by uncritically relying on Big Data analysis, public policy built upon these data sets will go against the directive of maximizing welfare on the grounds of the diminishing returns mentioned by Goodin.[202] The point is that based on welfare utilitarianism it also matters how the gains through efficiency are distributed. Implementing public policy that promotes the position of those that are already advantaged at the cost of marginalized groups of citizens would clearly violate any notion of maximizing welfare in a utilitarian calculation. On the grounds of the special responsibilities attributed to public servants by Goodin, public policy-makers are obligated to ensure that they correct for any such exclusions in the data sets.[203]

### 3.3 Scripted negative outcomes

As was seen in section 2.3, the employment of algorithms can make the interaction between public institutions and citizens more predictable for citizens, as well as, a more reliable process by standardizing much of it. This harkens back to the clarity provided by the rule-following nature of algorithms. Pasquale critiques that many of these rules seem, once released in real-world conditions, completely arbitrary.[204] With the FICO-score example, a person might unwittingly negatively influence his or her rating by lowering the overall amount of credit, because the algorithm takes into account the proportion of credit that is being used.[205] The internal logic of the algorithm creates a maze of arbitrary rules that individuals need to navigate through to avoid being penalized. For public institutions this maze needs to be navigated to gain access to the full array of goods and services. Given the arbitrary and counter-intuitive nature of some of these rules, being able to afford a guide is highly beneficial. According to O’Neil, only the affluent or well-connected will be able to afford the help needed to properly navigate this algorithmic maze.[206]

A second consideration is that algorithmic systems are opaque. Public institutions often work together with private companies. These private companies will be reluctant to give access to their algorithms, on the grounds of these algorithms being crucial for maintaining their competitive advantage. According to Pasquale, companies will are secretive about how

---

199 Lerman, “Big data and its exclusions.” 59.
200 Lerman, “Big data and its exclusions.” 59, 60.
201 Goodin, *Utilitarianism as a public philosophy*, 23.
202 Goodin, *Utilitarianism as a public philosophy*, 23.
203 G Goodin. “Public Service Utilitarianism as a Role Responsibility.” 322.
204 Pasquale, *The Black Box Society*, 22-25, 41.
205 Pasquale, *The Black Box Society*, 24
206 O'Neil, *Weapons of Math Destruction*, 61-64.

their algorithms are designed by claiming this is proprietary information.[207] For instance, the company that created the COMPAS sentencing algorithm does not disclose how it works based on these grounds.[208] Therefore, the COMPAS algorithm might be factoring in highly prejudicial information without the public, or the judges using the risk-score, being aware of it. Secondly, organizations necessarily need to limit the lucidity of how their algorithms function on the basis of protecting their systems against anyone trying to game the system. According to Kitchin, this is because knowledge about the design choices made during the development process allows individuals to play into the rules an algorithm is designed to follow.[209] For instance, once criminals learn the logic behind the COMPAS risk-scoring questionnaire they will be able to answer the questions in such a manner that they will be able to lower their risk-scores. Diakopoulos describes this effect as Goodhart's law.[210]

A third problem occurs when the same kind of algorithm or set of rules become ubiquitous. O'Neil shows that the proliferation of hiring algorithms leads to a cast of people becoming unemployable based on the widespread usage of the same algorithms in the hiring process.[211] Whereas in the past a person with a criminal record or mental illness might have been able to find an HR manager that would be willing to give that person a chance, now the HR manager would never be able to see or meet that person. The algorithm has already automatically cast every one aside with a mark beside their name. For the HR department, this is incredibly efficient, given that they no longer have to sort through large piles of job applications. The problem is that this standardization eliminates any chance specific groups of individuals have of escaping their situation.[212] Once the logic of an algorithm becomes ubiquitous it becomes inescapable. This is relevant for the interaction between public institutions and citizens, given that, if the public sector doubles down on conditions in the market, certain individuals will be harmed. This shows that at some points there are valid reasons to eschew from taking the most accurate or efficient route in order to avoid scripted negative outcomes that cause a certain group of citizens to be overwhelmingly disadvantaged.

### 3.4 Distinct errors

As was seen in section 2.4, the employment of algorithms will reduce the amount of errors rooted in human action. Automation will severely reduce the influence any given public servants have over the outcomes of a system. Furthermore, algorithmic scoring systems aid human decision-makers to avoid making mistakes. In many ways, it is the reliability and consistency of algorithms that ensure that human errors can be notably less common.

---

207 Pasquale, *The Black Box Society*,, 4, 96
208 Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner,"Machine Bias" Accessed June 21, 2018. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
209 Kitchin "Thinking critically about and researching algorithms" 20.
210 Diakopoulos, "Algorithmic-Accountability: the investigation of Black Boxes." 12.
211 O'Neil, *Weapons of Math Destruction*, 100.
212 O'Neil, *Weapons of Math Destruction*, 100.

However, specific kinds of algorithms are burdened with specific kinds of errors. For instance, Diakopoulos argues that classification algorithms will either produce false positives or false negatives.[213] This point also comes to the fore with the example of medical diagnosis software discussed in section 3.1. Kraemer et al. claim that it is unavoidable that the software used to classify whether or not a patient has a medical condition produce either false positives or false negatives, on the grounds of it being a judgment call where the threshold should be placed for when a patient is healthy. Whichever choice is made by the designer of the algorithm, certain cases will necessarily either be mistakenly be cast to the wayside or give rise to false alarms.[214] The point is that with classification algorithms it is impossible to fully avoid producing erroneous information, it is a matter subjective judgement with regards to how to balance the tradeoffs between the consequences of acting on false positives or false negatives. For public policy this can lead to resources being spend in an inefficient manner, due to the information falsely stating there is a problem that needs to be addressed. Contrariwise, if a threshold is set too high the algorithm will fail to recognize negative effects experienced by a groups of citizens. Policy-makers will subsequently be unable to address these issues or make incorrect decisions concerning the allotment of goods and services. Both these issues will restrict public servants in administering their duty to maximize public welfare.

Errors caused by the employment of algorithms and automation can severely hamper the maximizing of public interest on two grounds. Firstly, such errors can quickly have severe negative consequences for a person that is neigh on impossible to correct. Pasquale illustrates this with the example of a woman that was falsely flagged as a meth-dealer. This error in classification has led to her being turned down during job applications, as well as, her being unable to get access to credit to pay for important household repairs.[215] Resolving problems caused by the mistakes made by algorithms has proven to be exceedingly difficult. After the woman found out which data broker had made the mistake, an impressive feat in and of itself, she was told not much could be done, given that the data broker had already shared the erroneous information to a wide array of other data brokers. When data is being easily and quickly shared across many data sets controlled by a wide variety of organizations it is an impossible task to get any grips on what kind of information is stored where. In the public sector connecting databases might provide large gains in efficiency and allow for increased cooperation between public institutions. According to the Brenninkmeijer, it will also amplify the negative consequences experienced by citizens when one of these institutions makes a mistake. The ethereal nature of data makes it difficult to track how to erase classification mistakes or entirely be sure whether or not a problem is resolved.[216] The ease and speed data at which data can be shared further aggravates these problems.

[213] Diakopoulos, "Algorithmic-Accountability: the investigation of Black Boxes." 6.
[214] Kraemer, et al., "Is there an ethics of algorithms?" 254, 255.
[215] Pasquale, *The Black Box Society*, 33.
[216] Brenninkmeijer, "Meer dan een burger-servicenummer

The second reason that the employment of algorithms in public might impede maximizing public welfare is the widespread negative consequences that can be caused by automation. If an individual makes an incidental mistake this will be significantly different from an algorithmic system mass producing the same mistake. The effects of such automation in law enforcement is exemplified by the controversy surrounding the crippling fines awarded to Dutch citizens for not having a proper insurance for registered vehicles.[217] In 2011 a policy was introduced to automatically compare the database of registered cars against the list of insured vehicles, and send a fine to anyone who owned an uninsured registered vehicle.[218] The goal was to reduce the costs associated with the damages caused in accidents that involved uninsured cars. Problematic was that many of these registered vehicles had not been roadworthy for decades, and often already turned into scrap. The system churned out false positives on a massive scale considering that vehicles that are not roadworthy will not be involved in any accidents. A more fine-grained threshold would have been necessary to reduce the costs associated with uninsured vehicles while avoiding the severe negative consequences that followed citizens faced. The automation of the enforcement process led to the incarceration of nearly 20000 citizens.[219] This happened because public servant 404040, an automated classification system acting without human over-sight,[220] sent out swaths of fines that could only be appealed after a person managed to pay the fine. *De Nationale Ombudsman* points out that many were unable to do this. This barrier created an absurd and unjust situation where thousands of citizens got incarcerated over fines for vehicles they no longer owned or vehicles that had not been roadworthy for decades.[221] The report of *De Nationale Ombudsman* shows that it was those that were in precarious financial situations that suffered the brunt of the injustice caused by the false positives.[222] Brenninkmeijer argues that it is, therefore, vital to maintain an easily accessible pathway for citizens to appeal to a human authority or human decision-maker to quickly respond to the harm caused by automated systems.[223]

### 3.5 Human Context

As was seen in section 2.1, the employment of algorithms and Big Data analysis will create new opportunities for public institutions to administer their tasks. The ability to take proactive measures on the basis of faster data processing will enable public policy-makers to design more efficient and effective policies. However, these new opportunities can cause

---

[217] Frederiks, Jesse. “Het absurde Nederlandse boetebeleid: 18.733 celstraffen voor onverzekerde auto’s (die in de garage staan).” *De Correspondent,* May 4, 2017. https://decorrespondent.nl/6661/het-absurde-nederlandse-boetebeleid-18-733-celstraffen-voor-onverzekerde-autos-die-in-de-garage-staan/400734412639-88854651.
[218] Ministerie van Justitie en Veiligheid, M. V. J. (2017). 2585-Vermuldering WAM 30, 25.
[219] De Nationale Ombudsman,. “Gegijzeld door het systeem” Accessed June 21, 2018. https://www.nationaleombudsman.nl/system/files/bijlage/88985_Rapport_Gijzelingen_i-PDF.pdf, 13.
[220] Ministerie van Justitie en Veiligheid, M. V. J. (2017). 2585-Vermuldering WAM 30, 25.
[221] De Nationale Ombudsman, “Gegijzeld door het systeem” 2.
[222] De Nationale Ombudsman, “Gegijzeld door het systeem”19.
[223] Alex Brenninkmeijer, “Meer dan een burger-servicenummer”

significant collateral damage, most notably to marginalized groups. For instance, Brenninkmeijer, argues that a small mistake in the database of one public institution can cause years of grief for a person.[224] The speed and automation of the decisions made by algorithms amplify the negative consequences caused by such mistakes. From the example of the incarceration of thousands of Dutch citizens on the basis of fines bestowed for failing to get an insurance for non-existent cars, it becomes clear that in the context of the real-world there needs to be sufficient control measures to the influence of algorithmic systems.

These control measures need to at least compose of, as proposed by Brenninkmeijer giving citizens the ability to escape out of the grips of an algorithm by directly appealing to a human decision-maker that has sufficient authority and discretionary power to bend the rules. Everyone who ever found him or herself trapped in the loop of selection options of a helpdesk just to get to talk to a human person can surely see the merit in the proposal of Brenninkmeijer.[225] A second solution is leaving it up to the judges to correct the worst of the excesses. This would not meet the demands of a welfare utilitarian on the grounds of the possibility for many small infringements falling through the cracks, as well as, an inevitable delay-time before judges will be able to recognize that the law of the land no longer suffices. The third option is ensuring that during the design process human discretionary-power is built in. Front-loading such safeguards will necessarily limit the number of new options provided by the employment of algorithmic systems given that humans need to constantly be kept up to speed. However, it will also ensure that the most egregious cases of injustice can be corrected in a timely manner. Humans can use their expertise and experience to constantly steer Big Data-driven public policy towards outcomes that make sense in the context of the real-world, and they can be held be accountable to do so on the basis of their role responsibilities.

[224] Alex Brenninkmeijer, "Meer dan een burger-servicenummer"
[225] Alex Brenninkmeijer, "Meer dan een burger-servicenummer"

## 4. Framework for employment algorithms in the public sector

In this section, I will argue that the employment of algorithms should be assessed on the basis of the public welfare utilitarianism. Firstly, I will discuss the argument made by Goodin with regards to the merits of welfare utilitarianism as a normative guide for public action. Afterwards, I shall go into his notion of role responsibilities and how it can be used to argue for public servants having an obligation to maximize public welfare. Thirdly, I will combine the public welfare utilitarianism propagated by Goodin with the arguments made in sections 2 and 3 to create a framework for a normative assessment of algorithmic systems.

### 4.1 Utilitarianism and public policy

In *Utilitarianism as Public Philosophy* Goodin argues that utilitarianism is especially well suited to provide a moral underpinning for what the right course of action would be with regards to public affairs.[226] According to Goodin, many of the arguments used against utilitarianism as a normative guide for personal conduct do not apply in the realm of public affairs. Indeed, he even shows, most of the ailments of utilitarianism on a personal level are turned into advantages at the aggregate or public level.[227] For instance, the argument that utilitarianism can be considered to be impersonal or calculating may be failings for a normative guide on how to interact on a personal scale, however, when thinking about public policy decisions being detached or making impersonal calculations can be considered advantageous.[228] Goodin further argues that the abolishment of strict moral codes that need to be adhered to no matter the consequences is, for some, a troubling feature of utilitarianism. Goodin claims that for public officials breaking such moral codes is often a necessity in order to administer their duties correctly.[229] A public official might need to forsake telling the truth when being truthful would endanger public security or a diplomat might have to break a promise when it turns out that keeping his or her word would be against the interest of the people. Goodin adds to this that: "the public rather than private application of utilitarianism precepts helps use evade some of the most standard practical

[226] Goodin, *Utilitarianism as a public philosophy*, 4.
[227] Goodin, *Utilitarianism as a public philosophy*, 8-11.
[228] Goodin, *Utilitarianism as a public philosophy* 8-9.
[229] Goodin, *Utilitarianism as a public philosophy*, 10.

and practicality objections to the doctrine."[230]. The most notable of these is that it would be impossible to effectively calculate how much the maximizing of a good for one should be evaluated against the potential gains or losses of another person.[231] Goodin reflects this argument by stating that, in public debate, we already have to assume that "others are much like ourselves"[232], and that the same logic should be transferred when comparing added utility.[233]

This leaves open what needs to be maximized or compared in the utilitarian calculation. Goodin argues that the utilitarian maximand should be interests or general welfare.[234] According to Goodin, maximizing interests fits well with the making of public policy decisions due to the "relative commonality"[235] of what it means to further the interests of a person.[236] An added benefit of using public interests as a utilitarian maximand is that it promotes an equal and fair distribution of goods and services. More importantly, it can be used to make an argument for ensuring that the collateral damage caused by the implementation of public policy is not heaped on already disadvantaged groups. This can tip the scale in favor of protecting disadvantaged groups when assessing a policy that is likely to be otherwise producing gains through efficiency. These considerations are based on the observation made by Goodin that there are notable diminishing returns when trying to maximize interests.[237] At a certain point the added benefit of providing additional goods and services to already affluent or advantaged groups will barely register in the utilitarian calculation. Affluent or advantaged groups are likely to be able to supplement any gap in provided services by addressing personal means or their personal network. Likewise, making available goods and services to those that need them the most would rightfully be favored using such a utilitarian calculus. Goodin argues that such logic will "lead utilitarians to embrace policies and practices that are broadly egalitarian in form."[238]. This is especially relevant when addressing issues of bias or discrimination in algorithms, and to assess the effects of the employment of algorithms in public policy.

## 4.2 Egalitarian bend

Strikingly, there are barely any diminishing returns for the negative consequences experienced by already disadvantaged groups. Individuals that have experienced severe difficulties are likely to have depleted their reserves in order to stay afloat. Given that such reserves are likely to be limited, they will not be available to deal with the fall out caused by

---

[230] Goodin, *Utilitarianism as a public philosophy*, 18.
[231] Goodin, *Utilitarianism as a public philosophy*, 19,20.
[232] Goodin, *Utilitarianism as a public philosophy*, 20.
[233] Goodin, *Utilitarianism as a public philosophy*, 20, 21.
[234] Goodin, *Utilitarianism as a public philosophy*, 13.
[235] Goodin, *Utilitarianism as a public philosophy*, 14.
[236] Goodin, *Utilitarianism as a public philosophy*, 14.
[237] Goodin, *Utilitarianism as a public philosophy*, 23.
[238] Goodin, *Utilitarianism as a public philosophy*, 23.

public policy changes. When discussing whether the employment of an algorithm in the public sector adds or subtracts to public welfare a strong emphasis on the egalitarian bend is needed. Added to the notion of diminishing returns mentioned by Goodin is that the employment of algorithms has the potential to hit marginalized groups disproportionally hard. This will be because citizens that already find themselves in disadvantaged positions will likely not have access to the reserves needed to address the negative consequences of the employment of algorithms by public institutions. The car insurance example illustrates this point. Individuals that were heavily in debt did not have the funds to pay the fines, therefore, they were also unable to appeal the decision made by public servant 404040. Policy-makers should take this into account when designing public policy. In the next section, I will go into the foundation of such an obligation for public servants.

### 4.3 Role responsibilities for public servants

The strength of public welfare utilitarianism is that it gives an argument for the normative components behind practical matters in public policy. Practical concerns such as wastefulness or stability become morally relevant for the public sector based on the consequences the policy decisions have on maximizing welfare. For instance, I will argue in the following section, efficiency becomes an obligation with regards to public service on the grounds of it enabling more services to be provided for the public. This is relevant for the public sector given that citizens rely on public institutions for wide-ranging services, such as healthcare and social benefits. This reliance on public institutions, and by extension the tasks performed by public servants, make it pertinent to give a foundation for the obligation associated with public service.

Based on the discussion of utilitarianism as a public philosophy such an argument can specifically be provided for public servants. This would be built upon the notion of role responsibilities as discussed by Goodin in a later article. In *Public Service Utilitarianism as a Role Responsibility* Goodin shows how role responsibilities are related to public service, and how they sire certain duties and obligations. Following this argument a public servant merely needs to be reminded he or she is acting in an official capacity, to remind that public servant of the duties they have chosen to take on themselves by adopting the role of being a public servant.[239] According to Goodin: “Utilitarianism of some form or another is incumbent upon public policy-makers because of the peculiar tasks they face and because of the peculiar instruments available to them for pursuing those tasks”[240]. Goodin argues that public servants hold a special duty to further the interest of the public by virtue of the role they have taken upon themselves. He states that certain roles have specialized clusters of responsibilities attached.[241] For instance, a lifeguard on a beach has an obligation to be

[239] Goodin, “Public Service Utilitarianism as a Role Responsibility.” 322, 333.
[240] Goodin, “Public Service Utilitarianism as a Role Responsibility.” 320.
[241] Goodin, “Public Service Utilitarianism as a Role Responsibility.” 322.

vigilant, as well as, a duty not to fall asleep or wander off. This is based upon the reliance of the public on the lifeguard to intervene in case of emergency. Goodin calls such clusters of specialized responsibilities "role responsibilities"[242]. He further argues that these specialized responsibilities fit well within welfare utilitarianism, on the grounds of it being clear that the ability to adopt the obligations that go with them aid in maximizing welfare.[243] Public servants have these specialized clusters of responsibilities due to them acting in the capacity of public servants, and the reliance of others to perform the tasks associated with performing the role of public servant. An Archimedean point to answer the question why a public servant would have any reason to feel bound by any obligations tied to public services, it will be sufficient, according to Goodin, to remind that person he or she is acting in the context of public service.[244]

Based on the welfare utilitarianism propagated by Goodin in *Utilitarianism as Public Philosophy*, and the obligations that stem from role responsibilities, public policy-makers have a duty to design, implement, and steer public policy in such a way that it maximizes the added welfare for the general public. From this duty striving for efficiency in public policy receives a normative component. Given that the public relies on public sector institutions for a myriad of goods and services. This demand forces public servants to make difficult decisions on how to allocate scarce recourses. Poignantly, it are decisions made about the allocation of resources that are fundamentally public on the grounds of them being, to a large extent, garnered through taxation. Efficiency will aid in maximizing the total amount of goods and services a government is able to provide, based on each of tasks taking fewer recourses to be fulfilled. This corresponds with the maximand proposed by Goodin.[245] As was argued for in section 2.2, the employment of algorithms will lead to such a growth in efficiency that it would be irresponsible for public policy-makers to forsake in fully exploring these options

### 4.4 Efficiency as a moral duty

There is a distinct normative aspect to maximizing efficiency in the public policy. An argument for this can be broadly based on the welfare utilitarianism propagated by Goodin.[246] Furthermore, the notion of role responsibilities described by Goodin can be used to serve as the foundation for the moral obligation of public servants to fully explore the possible gains through efficiency that can be procured from the employment of algorithms. Given that the public relies on public instructions to provide fundamental goods and services, and the finite nature of the number of recourses available to public institutions, gains through efficiency will allow for public institutions to be better able to further general

[242] Goodin, "Public Service Utilitarianism as a Role Responsibility." 322
[243] Goodin, "Public Service Utilitarianism as a Role Responsibility." 320, 321, 322.
[244] Goodin, "Public Service Utilitarianism as a Role Responsibility." 333.
[245] Goodin, *Utilitarianism as a public philosophy*, 13.
[246] Goodin, *Utilitarianism as a public philosophy,* 26.

public welfare. Most notably, this will be because a more efficient use of resources in one area will lead to more being resources being available for another. This will cause public institutions to be able to provide a higher amount of overall goods and services or to mitigate the negative effects of budget cuts. According to Goodin, public servants and public policy-makers have the responsibility to maximize the interests of the public. This responsibility obligates public servants to act in a manner that furthers public welfare. Whichever way this is to be defined, maximizing it will require making efficient use of scarce public resources.[247]

### 4.5 Framework for using algorithms in the public sector

The employment of algorithmic systems can lead to vast improvements in how public institutions perform their tasks in for five reasons. Firstly, public policy-makers gain the ability to take a proactive stance when trying to increase the welfare of the public on the basis of the necessary information becoming available significantly faster through algorithmic data mining. Secondly, gains through efficiency can be realized through automation and improved allocation of recourses on the basis of superior information. Gains in efficiency can subsequently be translated into welfare maximization based on more services being provided to the public. Thirdly, automation and risk-scores will make the implementation of public policy more consistent and reliable. Fourthly, the negative consequences caused by human errors can be curtailed through automation, and algorithms can be used as a counterweight for human biases in the decision-making process. Fifthly, algorithms introduce a form of objectivity that will lead to public servants making better decisions. All of these combined make the employment of algorithms quintessential for public servants to fulfill their obligation to serve the public interest.

This is not to say that there are no limits to the benefits afforded by the use of algorithms and Big Data Analytics in the public sector. Each of the five advantages associated with the employment of algorithms has a drop-off point. Firstly, algorithms cannot entirely break free from human bias due to the reliance of algorithms on data sets. If these data sets are filled with prejudice, the algorithm will be stuck in a feedback loop that will have an adverse impact on public welfare. Secondly, there are various drawbacks to optimizing for efficiency. At a certain point, efficiency leads to inflexibility. Public institutions need this flexibility due to the ever-changing problems they face, and the varied nature of the citizenry. Therefore, public institutions need to maintain some degree of inefficiency, in order to be able to maximize public welfare.  A third restriction to the extent public institutions ought to rely on algorithms is that once certain algorithmic systems become ubiquitous they will disparately place the burdens on specific groups. On the grounds of a lack of diminishing returns with regards to these adverse effects, public servants are obligated to ensure that these are not

[247] Goodin, "Public Service Utilitarianism as a Role Responsibility." 322.

heaped on to already disadvantaged groups. Fourthly, algorithms are prone to make certain types of errors. A human decision-maker needs to have some control over the process to mitigate the consequences of such errors. This will create a necessary constraint on how far automation can be pushed in public policy. Finally, given the far-reaching consequences of the actions made public institutions there indubitably needs to be a way to appeal to a human public-servant with the discretionary power to bend the rules, and free a person out of the grips of an algorithmic system. This will prove to be a constant, and worthwhile, constraint on the utilization of algorithmic systems in the public sector.

All of the arguments discussed thus far should be seen in the light of a utilitarian calculation. Specifically, whether or not it leads to the maximization of public interest. This showcases the merit of utilitarianism as a public philosophy. Even more so if the egalitarian bend in welfare utilitarianism is sufficiently emphasized to ensure that not one specific group of citizens should have to face the brunt of societal injustice. The strength of Goodin’s notion of role responsibilities is that it provides a normative foundation for the obligations of public servants. This allows for welfare utilitarianism as a public philosophy to both support an argument for efficiency in public policy, and against discriminatory outcomes of the actions performed by public institutions. Ultimately, public policy-makers will need to balance the gains and disadvantages caused by relying on algorithmic systems. In the next section, this will be done for predictive policing.

## 5. Predictive Policing

In this section, I will use predictive policing  as a case study to showcase how the ethical framework can be applied to assess how the employment of a specific algorithm adds or subtracts to the maximization of public welfare. Predictive policing is the use of statistical analysis to make informed decisions on how to allocate police resources. Ensign et al. describe predictive policing as: “Given historical crime incident data for a collection of regions, decide how to allocate patrol officers to areas to detect crime.”[248] They specify that the desired outcome of introducing predictive policing is the prevention of crime.[249] I will use two parts of the ethical framework discussed in section 4 to show the merits and demerits of predictive policing. In section 5.1, I shall argue that predictive policing ought to be implemented on the basis of the significant gains through efficiency it can provide. According to Mali et al., the implementation of CAS allows police officials to be proactive in the allocation of recourses.[250] More efficient and effective allocation of recourse will, in turn, enable police departments to provide for more and better services across the board. A limitation will be that some amount of leeway needs to be maintained for police officers to be able to properly fulfill their roles as public servants.  During their study  Mali et al. discovered that at some points the discretionary power of individual police officers will be in conflict with the implementation of CAS.[251] Given the importance of flexibility in how public institutions approach citizens, and the severity of the consequences police action can have on the lives of individuals, police officers need to be awarded sufficient leeway to make their own decisions when they interact with citizens. This will aid in maximizing public welfare because it can act as a safeguard against the worst negative side effects of the employment of algorithmic systems in the public sector.

### 5.1 Efficiency and Inflexibility

Predictive policing will aid in maximizing public welfare on the grounds of the gains that can be made through efficiency. Mali et al. argue that implementing a predictive policing system is vital for police departments to be able to administer their duties in the future.[252] CAS

[248] Danielle Ensign, Sorelle Friedler, Scott A., Neville, et al.“ Runaway feedback loops in predictive policing“ *Proceedings of Machine Learning Research*, 81, (2018),  2.
[249] Ensign et al, Runaway feedback loops in predictive policing“ 1,2.
[250] Mali, Bronkhorst-Griesen and Den Hengst, “Predictive policing: lessen voor de toekomst“  57, 58.
[251] Mali, Bronkhorst-Griesen and Den Hengst, “Predictive policing: lessen voor de toekomst“  71.
[252] Mali, Bronkhorst-Griesen and Den Hengst, “Predictive policing: lessen voor de toekomst“  57-58.

involves using algorithms to automatically analyze a wide array of data to give a prediction where it is most likely that criminal behavior will occur. The result is a color-coded map that shows at which locations police patrols will be most effective in preventing crime or catching criminals. This information will enable police officials to preemptively allocate extra resources to specific locations, rather than having to react after a crime has occurred. They argue this will become increasingly relevant considering that a reactive stance is no longer enough to solve the problems police officials currently face.[253] To illustrate the importance of being able to take a proactive stance Mali et al. give an example of designing a policy to prevent people from drowning. When the objective is to save people from drowning in a river, and the amount of people at risk of drowning grows to such a degree that it exceeds the available recourses, it is necessary to find a solution that addresses the underlying root of the problem. In other words, instead of reacting to every person that is at risk of drowning, the goal should be to find a solution to why so many people fall in the river or are unable to swim.[254]

Police departments will be especially more effective in preventing burglaries when implementing predictive policing. From a study about burglaries it is shown that they tend to happen in clusters, with burglars hitting the same area multiple times in a short span of time.[255] This enables the making of highly accurate and specific predictions. Mali et al. conclude that acting on such predictions has led to a notable reduction in burglaries, as well as, overall crime.[256] Perry points out that a map based on these predictions cannot be optimized for accuracy alone on the grounds of specificity being a necessary requirement for police officials to be able to make effective decisions regarding where to send police patrols.[257] CAS meets both these demands on the basis of the hexes being sufficiently specific, 125m by 125m, while still being highly accurate.[258] The main benefit of such a proactive stance is that crime will be prevented due to police presence acting as a deterrence for criminal activity.[259] Mali et al. state that an additional advantage is that police patrols will be more likely to catch criminals in the act.[260]

This is pertinent for maximizing public welfare on two grounds. Firstly, better informed decisions regarding where to send police patrols will prevent public resources being wasted. Mali et al. state that predictive policing enables police officers to be at the right place at the right time.[261] Consequently, the ability to specifically target certain areas means that police officials can forego having to cast a wide-net when assigning police patrols. This in turn

---

[253] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst"  16.
[254] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst"  57.
[255] G. O Mohler, M. Short,. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. "Self-exciting point process modeling of crime." *Journal of the American Statistical Association*, 106 (493), (2011): 100-108.
[256] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst" 30.
[257] Perry, "Predictive Policing" 23.
[258] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst"  30, 31.
[259] Moses L. Bennet and Janet Chan, "Algorithmic prediction in policing: assumptions, evaluation, and accountability." *Policing and Society*, (2016): 5.
[260] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst" 30.
[261] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst" 69.

implies those resources are available for securing large events or to mitigate the negative consequences of personnel shortages. The point is that if police officials can use the predictions to make better informed decisions with regards to the allocation of recourses, police departments will have more resources to provide other important services to that aid in maximizing public welfare. Based on this, there is an obligation to explore the potential merits of a predictive policing model. Secondly, preventing crime will lead to a lower amount of damages and costs associated with the aftermath of criminal behavior. This includes the costs to the wellbeing of citizens that are the victims of crimes, as well as, having more resources available to solve past crimes. Assuming that some of these past crimes have been perpetrated by repeat-offenders, solving them will cause the prevention of further crimes, thus lowering the overall crime-rate. The point is that predictive policing enables more effective and efficient allocation of police recourses. Most notably with regards to specific crimes that follow highly predictable patterns, such as burglary.[262] Consequently, police officials will have more resources available to face other challenges, such as securing large events or crime-solving. Given this potential, and the role responsibilities enjoyed by public policy-makers, the implementation of CAS should be further explored.

There are three possible limitations to the gains made through efficiency. Mali et al. show that police officers require a high amount of discretionary power in order to administer their duties. The implementation of predictive policing has proven to be in conflict with this freedom enjoyed by police officers. According to Mali et al., it can cause tensions to arise between police officials and police officers.[263] The discretionary power needed for police officers to function will serve as a constraint on how far predictive policing can change police policy. Police officers need to be given sufficient leeway in how they give substance to the policing required to make predictive policing effective. Perry argues that predictions need to be coupled with the work of experienced and capable police officers in order to produce actual results.[264] Between the work done by the police officers and the predictive models used by police officials there might be a disconnect. It is up to police officers and police officials to bridge the divide between the prediction and the actual policing. Police officers will need to be given sufficient leeway to do this, even if this comes at the cost of efficiency.

From the example of the car insurance illustrates how important such leeway can be for an organization to avoid unfairly burdening already disadvantaged groups. Based on the egalitarian bend in public welfare utilitarianism public policy-makers are obligated to avoid already disadvantaged groups facing the brunt of the negative consequences that are a byproduct of otherwise beneficial public policies.[265] Especially considering that the rigidity of a bureaucratic system might combine poorly with algorithmic systems once they start running too efficiently. With the car insurance example, it was the standard practice to incarcerate citizens that are unwilling or unable to pay their fines, and the impossibility to

---

[262] Mohler et al., “Self-exciting point process modeling of crime” 100.
[263] Mali, Bronkhorst-Griesen and Den Hengst, “Predictive policing: lessen voor de toekomst“ 71, 72, 73.
[264] Perry, *Predictive Policing*, 28.
[265] Goodin, *Utilitarianism as a public philosophy* 23.

appeal unjust fines without first paying those fines, that caused a system that was strikingly detrimental to already disadvantaged groups.[266] Giving police officers sufficient leeway and awareness will avoid much of these problems. This is of vital importance for the ability of predictive policing to aid in maximizing public welfare, because a similar problem might occur as a byproduct of widespread implementation of predictive policing.

Given that police officials will send more police patrols to specific areas this concentration of police resources will have two negative side-effects. The first are feedback-loops, these will be discussed in the following section, and the second is that it constitutes as an extra tax on minor infractions in for the residents of those areas. A person living in an area with additional police patrols will be far more likely to receive a fine for minor infractions that are quite normal for citizens to commit in their daily lives. Additional police presence in an area will cause more of these small misdemeanors to be caught, and therefore, anyone living in those areas needs to either shore up his or her behavior or pay what is a de facto added tax for living at these algorithmically targeted locations. The problem with fines and groups in precarious financial situations can be taken from the car insurance example discussed in section 3.4. The solution is by maintaining sufficient discretionary power on the part of the police officer. This should be done by giving police officers leeway to desist from giving a fine, as well as, to forego on any policy that severely impedes on this discretionary power of polices officers. An example of such a policy would be introducing a quota with regards to the minimum amount of fines any individual police officer needs to produce. Such a policy limits the discretionary power of police officers on the grounds of police officers feeling that it might force their hand when they have to decide whether or not they should give a person a fine. It also reinforces any potential feedback-loops. The latter problem will be discussed in the following section.

## 5.2 Objectivity and Bias in algorithms

Boosters of predictive policing praise it to be a more objective form of policing. A police commander of the LAPD makes this exact argument.[267] Mali et al. argue that what is innovative about predictive policing is that it replaces the old heuristics used by police officers with statistics and predictions.[268] Subsequently, the accuracy of the predictions allows police officials to make better informed decisions with the necessary amount of confidence that it can be considered responsible to act on the information provided by the algorithmic predictions. This is relevant given how important the prevention and solving of crime is for public wellbeing. However, relying on data-driven models is, as was seen in

---

[266] De Nationale Ombudsman,. "Gegijzeld door het systeem"48.

[267] Justin Jouvenal, "Police are using software to predict crime. Is it a 'holy grail' or biased against minorities?" *The Washington Post*, November 17, 2016. https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8_story.html?noredirect=on&utm_term=.6da65797fca8.

[268] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst" 59.

section 3, plagued with ethical problems. The most notable of them for predictive policing is the reliance on historical data. With CAS the predictions rely heavily on the data derived from past police action.[269] This makes the content of these data sets, and how it came about, relevant for a normative study, on the basis of it being pertinent for understanding how the algorithmic system functions.

Problematic is that several studies show that in the historical crime data are rife with bias and prejudice. For instance, According to Kristian Lum and William Isaac, historical crime data are influenced by subjective criteria such as the bias of those making the arrests or the willingness of groups to report a crime.[270] After feeding these data sets to a predictive policing algorithm widely used in the U.S., PredPol, they found out that the PredPol models are a reflection of the biases that are present in these historical crime data sets.[271] This is not to say that CAS will necessarily run into the same complications. If the historical crime data used for CAS is devoid of any subjective factors or the algorithm used and developed for CAS corrects for any such biases, the predictions furnished by the CAS algorithm will be unbiased. However, based on the point made by Kitchin, that an algorithm is always being updated, it is possible that such bias can creep into the algorithm at a later point.[272]

Added to this is the possibility that police departments that are using predictive policing systems might unwittingly launder biased data into objective data sets. This is the mechanism Pasquale described, seen in section 3.1, with regards to loan scores and how they target minority groups. When companies, on the basis of biased data, ask higher interests rates from minorities this will lead to members of those groups to be more likely to default on their loans, thereby proving that the data was correct. Pasquale argues this creates a self-reinforcing system that spawns seemingly objective data from subjective historical data.[273]

According to Crawford and Schultz: "the predictions that these policing algorithms make - that particular geographic areas are more likely to have crime- will surely produce more arrests in those areas by directing police to patrol them. This, in tum, will generate more 'historical crime data' for those areas and increase the likelihood of patrols."[274]. The problem is that introducing a predictive policing policy places a heavy burden on the crime data collected by law enforcement in the past. Lum and Isaacs have shown how this might be the Achilles heel of predictive policing systems.[275] The problem is even more pernicious than described by Crawford and Schultz. Considering that the algorithmic systems are sold as being objective and value neutral, police officers might acquire a trained-bias towards the citizens living in the locations they are send to the most. They will learn to associate certain

---

[269] Mali, Bronkhorst-Griesen and Den Hengst, "Predictive policing: lessen voor de toekomst" 39.
[270] Kristian Lum and William Isaac, "To predict and serve?." *Significance*, *13*(5), (2016): 15.
[271] Lum and Isaac "To predict and serve?" 18.
[272] Kitchin "Thinking critically about and researching algorithms" 18.
[273] Pasquale, *The Black Box Society*, 41.
[274] Crawford & Schultz," Big data and due process" 103, 104.
[275] Lum and Isaac "To predict and serve?" 5, 18.

areas, and the groups living there, with crime. This might feed into already existing prejudices, and it will at some translate into new crime data. This new data will further accelerate the feedback-loop, and cause a similar self-reaffirming loop with the pre-existing prejudices present for an individual police officer.

For public policy-makers it is exceedingly difficult to gauge the effectiveness of predictive policing. The feedback-loop will make it so that the algorithm keeps justifying its own predictions. Crucially, the predictions do not exist in a vacuum. From the point made by Perry predictive policing can only be effective once it is translated into police action.[276] These actions create their own data points, and the situations described by Crawford and Schultz arises.[277] From the perspective of the policy-makers it is therefore important to look for other sources of information about the success of predictive policing.

Once predictive policing systems get caught in a feedback-loop they are able to cause clearly unjust results. This negates any gains that might be had on the basis of efficiency. Efficiency should not be seen as a goal in and of itself. The point is to add to maximizing public welfare. Once an algorithmic system goes rampant public policy-makers ought to find ways to repair the situation or make the system work. For predictive policing the potential benefits are too great to leave on the table. Policy-makers will therefore be tasked to find the correct safeguards  to cure the system of its ailments without sacrificing the benefits.

---

[276] Perry, *Predictive Policing*, 28.

[277] Crawford & Schultz,“ Big data and due process” 103, 104.

## Conclusion

In this paper the purpose was to provide an outline for an ethical framework to assess the employment of algorithms in the public sector. This has to be done by identifying the relevant aspects of the use of algorithms by public institutions, and combine this with a supporting ethical theory that gives the framework a normative structure. The former necessitates studying the most relevant arguments about the merits and demerits of employing algorithms.

In the second section, several of the arguments for the use of algorithms were discussed. Most importantly, the employment of algorithms will lead to significant gains through efficiency. As was seen from the U.S. Veteran Affairs example mentioned by Joseph and Johnson in section 2.2, automation can make the interaction between public institutions and citizens significantly faster and more efficient.[278] Added to this is that faster information processing will create new policy options for public policy-makers. This was discussed in section 2.1 with the example of the use of Twitter data to create predictions about influenza outbreaks.[279] In section 2.3, the relevance of algorithmic scoring systems was considered. For instance, risk-scores can be used to aid judges to ensure subjective criteria do not play a part in their sentencing.[280] Algorithms can also be used to mitigate the risk of human error. In section 2.4, I discussed the example of medical diagnosis software being able to quickly and reliably diagnose diseases.[281] It is the ability of algorithms to quickly and reliably process large data sets that will provide a source of information that can mitigate the risk of human errors. In 2.5, I discussed how algorithms are considered to be a source of objectivity. This is relevant for the case study of predictive policing, considering that the users of the systems believe the that the algorithms produce objective results.[282]

The argument for the employment of algorithms in the public sector derives a normative component on the basis of how it influences the lives of citizens. According to Robert Goodin, public policy should be judged along the lines of how it fares in maximizing public

[278] Joseph and Johnson, "Big data and transformational government." 45, 46.
[279] Achrekar et al., "Predicting flu trends using twitter data" 716, 717, 718.
[280] Kehl et al., "Algorithms in the Criminal Justice System" 11.
[281] Naik and Bhide, "Will the future of knowledge work automation transform personalized medicine?" 51
[282] Jouvenal, Justin. "Police are using software to predict crime. Is it a 'holy grail' or biased against minorities?" *The Washington Post*, November 17, 2016. https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8_story.html?noredirect=on&utm_term=.6da65797fca8.

welfare. The arguments above each show the employment of algorithms can aid in maximizing public welfare. This makes them ethically relevant for assessing the success of the implementation of an algorithm in the public sector. A further ethical component is that public policy-makers are obligated to explore these options to maximize public welfare. Goodin argues that public servants have an obligation to aid in maximizing public interest due to the responsibilities associated with public service. Taking on the role of public servant entails adopting the responsibilities that come with this role. Added to this is an obligation for public policy-makers to avoid algorithms that subtract from maximizing public welfare.

In the third section, I discussed several aspects that illustrate how the employment of an algorithmic system can negatively impact the maximization of public welfare. The main concerns are that algorithms can inherit past biases from the data sets that are fed to them. This proved to be relevant for the case study considering Kristian Lum and William Isaac discovered that the predictions provided by a predictive policing algorithm reflected the past biases that were found in the historical crime data.[283] In section 3.2, I discussed possible limitations to the gains made through efficiency. According to Casey Haskins, an organization can become so efficient that it becomes inflexible.[284] This is relevant considering that public institutions will be required to provide wide-ranging services to citizens. Inflexibility will impair public institutions to provide these services in a proper manner. Another relevant aspect is that algorithms are capable of making mistakes. Especially, the mistakes made by classification algorithms can lead to significant negative consequences for a person. The car example showcases how a poorly designed classification algorithm employed by a public institution severely subtracts from the maximization of public welfare.[285]

The main argument is that when these interlocking arguments are combined with the welfare utilitarianism propagated by Goodin an ethical framework can be devised that can be applied to a specific case study. Applying the framework will highlight what the ethically relevant aspects of the employment of that specific algorithm are, and how they relate to maximizing public welfare. Based on the notion of role responsibilities of public servants these considerations can then be used as a foundation for the obligation of public policy-makers to either further implement the algorithm, refrain from doing so, or build in the necessary safeguards. In this paper two legs of the ethical framework were applied to the predictive policing algorithms. From this it was apparent that the potential gains in efficiency are sufficiently great that public policy-makers are obligated to look for the most effective way to use this new technology. However, much of these gains stand to be negated if the predictive policing algorithms churns out biased models that lead to the feedback-loops described by Kate Crawford and Jason Schultz.[286] Efficiency cannot be seen as goal in and of itself. Although, efficiency gains a moral component when designing public policy on the basis of the welfare added by the additional services that can be provided, it cannot weigh

[283] Lum and Isaac, "To predict and serve?" 15, 18.
[284] Haskins, "The drawbacks of efficiency"
[285] De Nationale Ombudsman,. "Gegijzeld door het systeem" 19.
[286] Crawford & Schultz," Big data and due process" 103, 104

up to the negative consequences such a feedback-loop can have once it targets areas with already marginalized groups. Especially, considering the egalitarian bend of public welfare utilitarianism public policy-makers are obligated to endeavor to resolve such feedback-loops.

Considering that feedback-loops justify their own predictions it will be difficult for public policy-makers to notice them. This opens up new areas for research based on solving the problem of how to detect and resolve these feedback-loops. Especially, once systems start to become faster and more complex it will be difficult to resolve complications that arise from the use of algorithms. Further study can be done about the ethical aspects of solutions of the complications associated with the employment of algorithms. For instance, how to balance efficiency and accuracy against these solutions.

## Bibliography

Achrekar, Harshavardhan, Avinash Gandhe, Ross Lazarus, Ssu-Hsin Yu, and Benyuan Liu. "Predicting flu trends using twitter data." In *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, pp. 702-707. IEEE, 2011.

Anderberg, Rozita H., Caroline Hansson, Maya Fenander, Jennifer E. Richard, Suzanne L. Dickson, Hans Nissbrandt, Filip Bergquist, and Karolina P. Skibicka. "The stomach-derived hormone ghrelin increases impulsive behavior." *Neuropsychopharmacology* 41, no. 5 (2016): 1199.

Ali, Tariq. "Will programming be automated? (A Slack Chat and Commentary)" Accessed June 21, 2018. https://dev.to/tra/will-programming-be-automated-a-slack-chat.

Angwin, Julia, Larson, Jeff, Mattu, Surya and Kirchner, Lauren. "Machine Bias" Accessed June 21, 2018. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Barocas, Solon and Selbst, Andrew D., "Big data's disparate impact." *California Law Review,* 104, (2016): 671-732, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899.

Barret, Nick. "Hostile environment: The dark side of nudge theory." Accessed June 21, 2018. http://www.politics.co.uk/comment-analysis/2018/05/01/hostile-environment-the-dark-side-of-nudge-theory

Bennett Moses L. and Chan, Janet. "Algorithmic prediction in policing: assumptions, evaluation, and accountability." *Policing and Society*, (2016): 1-17, https://www.tandfonline.com/doi/pdf/10.1080/10439463.2016.1253695?needAccess=true

Blair, Irene V., Charles M. Judd, and Kristine M. Chapleau. "The influence of Afrocentric facial features in criminal sentencing." *Psychological science* 15, no. 10, (2004): 674-679.

Bomholt, James. "Program Synthesis Explained." Accessed June 20, 2018. https://homes.cs.washington.edu/~bornholt/post/synthesis-explained.html

Boyd, Danah, and Kate Crawford. "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon." *Information, communication & society* 15, no. 5, (2012): 662-679.

Brenninmeijer, Alex. "Meer dan een burger-servicenummer." Accessed June 20, 2018. https://www.nationaleombudsman.nl/uploads/tijdschrift_voor_conflicthantering_hoeveel_legitimatie_schept_lokale_democratie.pdf

Bozdag, Engin."Bias in algorithmic filtering and personalization." *Ethics and information technology*, 15 no. 3, (2013): 209-227, https://link.springer.com/article/10.1007/s10676-013-9321-6.

Crawford, Kate, and Jason Schultz. "Big data and due process: Toward a framework to redress predictive privacy harms." *BCL Rev.* 55 (2014): 93.

Danaher, John. "The threat of algocracy: Reality, resistance and accommodation." *Philosophy & Technology* 29, no. 3 (2016): 245-268.

Dark Patterns “What are Dark Patterns?” Accessed June 18 2018. https://darkpatterns.org/

De Nationale Ombudsman,. “Gegijzeld door het systeem” Accessed June 21, 2018. https://www.nationaleombudsman.nl/system/files/bijlage/88985_Rapport_Gijzelingen_i-PDF.pdf

Diakopoulos, Nicholas. "Algorithmic-Accountability: the investigation of Black Boxes." *Tow Center for Digital Journalism* (2014).

Englich, Birte, Thomas Mussweiler, and Fritz Strack. "Playing dice with criminal sentences: The influence of irrelevant anchors on experts’ judicial decision making." *Personality and Social Psychology Bulletin* 32, no. 2 (2006): 188-200.

Ensign, Danielle, Sorelle, Friedler A., Neville, Scott et al.“ Runaway feedback loops in predictive policing“ *Proceedings of Machine Learning Research*, 81, (2018): 1-12, https://arxiv.org/pdf/1706.09847.pdf.

Fogg, Kathleen. “What Predicts Loan Repayment at Auto Capital?”, Honors Project, Spring, (2016): 1-28. http://commons.lib.jmu.edu/cgi/viewcontent.cgi?article=1240&context=honors201019

Frederiks, Jesse. “Het absurde Nederlandse boetebeleid: 18.733 celstraffen voor onverzekerde auto’s (die in de garage staan).” *De Correspondent,* May 4, 2017. https://decorrespondent.nl/6661/het-absurde-nederlandse-boetebeleid-18-733-celstraffen-voor-onverzekerde-autos-die-in-de-garage-staan/400734412639-88854651.

Friedman, Batya, and Nissenbaum, Helen. "Bias in computer systems." *ACM Transactions on Information Systems (TOIS),*14, no. 3 (1996): 330-347. https://dl.acm.org/citation.cfm?id=230561.

Gillespie, Tarleton. "The relevance of algorithms." *Media technologies: Essays on communication, materiality, and society*, 167, (2014): 167-199, http://culturedigitally.org/2012/11/the-relevance-of-algorithms/.

Gillespie, Tarleton. "Can an Algorithm be Wrong?." *Limn*, *1*(2), (2012): 16-20. https://limn.it/articles/can-an-algorithm-be-wrong/.

Goodin, Robert, E. “Public Service Utilitarianism as a Role Responsibility.” *Utilitas*, 10(3), (1998): 320-336. https://www.cambridge.org/core/journals/utilitas/article/public-service-utilitarianism-as-a-role-responsibility/D82CDFAD2C7B784AD0FA14328F57D73E

Goodin, Robert. E. *Utilitarianism as a public philosophy*. New York: Cambridge University Press, 1995.

Hand, David J. "Principles of data mining." *Drug safety* 30, no. 7 (2007): 621-622.

Haskins, Casey P. “The drawbacks of efficiency?” *The Chautauquan Daily*, August 16, 2011. https://chqdaily.wordpress.com/2011/08/16/the-drawbacks-of-efficiency/.

Hill, Robin K. "What an algorithm is." *Philosophy & Technology*29, no. 1, (2016): 35-59. https://link.springer.com/article/10.1007/s13347-014-0184-5.

Höchtl, Johann, Peter Parycek, and Ralph Schöllhammer. "Big data in the policy cycle: Policy decision making in the digital era." *Journal of Organizational Computing and Electronic Commerce* 26, no. 1-2 (2016): 147-169.

Jacobson, Ralph. “2.5 quintillion bytes of data created every day. How does CPG & Retail manage it?” *IBM*, April 24, 2013. https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/.

Joseph, Rhoda C., and Norman A. Johnson. "Big data and transformational government." *IT Professional* 15, no. 6 (2013): 43-48.

Jouvenal, Justin. “Police are using software to predict crime. Is it a ‘holy grail’ or biased against minorities?” *The Washington Post*, November 17, 2016. https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8_story.html?noredirect=on&utm_term=.6da65797fca8.

Kehl, Danielle, L., Guo, Priscilla and Kessler, Samuel. "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing", *Berkman Klein Center for Internet & Society*, (2017): 1-36, https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf?sequence=1.

Kim, Gang-Hoon, Silvana Trimi, and Ji-Hyong Chung. "Big-data applications in the government sector." *Communications of the ACM* 57, no. 3 (2014): 78-85.

Kitchin, Rob. "Thinking critically about and researching algorithms." *Information, Communication & Society* 20, no. 1 (2017): 14-29.

Kraemer, Felicitas, Kees Van Overveld, and Martin Peterson. "Is there an ethics of algorithms?." *Ethics and Information Technology* 13, no. 3 (2011): 251-260.

Lerman, Jonas. "Big data and its exclusions." *Stanford Law Review*, Online, 66, (2013): 55-63. https://www.stanfordlawreview.org/online/privacy-and-big-data-big-data-and-its-exclusions/.

Levitin, Anany. *Introduction to the design & analysis of algorithms*. Boston: Pearson, 2012.

Lorch, Donatella. "Openly Gay in Blue: Officers Tread Warily."*The New York* Times, July 13, 1992. https://www.nytimes.com/1992/07/13/nyregion/openly-gay-in-blue-officers-tread-warily.html.

Lum, Kristian and Isaac, William. (2016). "To predict and serve?." *Significance*, *13*(5), (2016): 14-19. https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x.

MacBride, Elizabeth. "How AI Aids Small Business Hiring: An Interview With ZipRecruiter's CEO." *Forbes*, October 31, 2017. https://www.forbes.com/sites/elizabethmacbride/2017/10/31/meet-the-jobs-startup-with-leverage-to-bring-google-and-facebook-to-the-table/2/#f6543d3235f7.

Mali, Bas, Carla Bronkhosrt-Giesen, en Marielle Den Hengst."Predictive policing: lessen voor de toekomst) 2017. Accessed June 20, 2018. https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/93263.PDF.

Mayer-Schönberger, Viktor and Cukier, Kenneth. *Big Data*. New York: Houghton Mifflin Harcourt Publishing Company, 2011.

Ministerie van Justitie en Veiligheid, M. V. J. (2017). 2585-Vermuldering WAM 30, 1-94.?

Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. "The ethics of algorithms: Mapping the debate." *Big Data & Society* 3, no. 2 (2016): 1-21.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. "Self-exciting point process modeling of crime." *Journal of the American Statistical Association*, 106 (493), (2011): 100-108, https://www.tandfonline.com/doi/abs/10.1198/jasa.2011.ap09546.

Naik, Gauri, and Sanika S. Bhide. "Will the future of knowledge work automation transform personalized medicine?." *Applied & translational genomics* 3, no. 3 (2014): 50-53.

O'Neil, Cathy. "Interview with Cathy O'Neil, author, Weapons of Math Destruction" , *Beyond the bookcast* , November 14, 2016. Accessed June 19, 2018. http://beyondthebookcast.com/wp-content/uploads/2016/11/ONeilTranscript.pdf.

O'Neil, Cathy. *Weapons Of Math Destruction.* London: Penguin Books, 2011.

Pasquale, Frank. *The Black Box Society.* Cambridge, Massachusetts: Harvard University Press, 2016.

Perry, Walt L. *Predictive policing: The role of crime forecasting in law enforcement operations*. Rand Corporation, 2013.

Pilcher, June J., and Allen I. Huffcutt. "Effects of sleep deprivation on performance: a meta-analysis." *Sleep* 19, no. 4 (1996): 318-326.

Rieder, Gernot, and Judith Simon. "Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data." *Big Data & Society* 3, no. 1 (2016): 1-6.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. "Can an Algorithm be Unethical?." *Ann Arbor*: 1-33.

The Interaction Design Foundation. "User Interface (UI) Design", *The Interaction Design Foundation*. Accessed June 19, 2018.https://www.interaction-design.org/literature/topics/ui-design

UIE. "Dark Patterns and the Ethics of Design" *Adventures in UX design*, 19, November 29, 2017. https://medium.com/adventures-in-ux-design/dark-patterns-and-the-ethics-of-design-31853436176b.

Vedder, Anton, and Laurens Naudts. "Accountability for the use of algorithms in a big data environment." *International Review of Law, Computers & Technology* 31, no. 2 (2017): 206-224.

Wu, Xindong, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan et al. "Top 10 algorithms in data mining." *Knowledge and information systems* 14, no. 1 (2008): 1-37.

Yiu, Chris. "The big data opportunity." *Policy exchange* 8 (2012): 1-36.

Zarsky, Tal Z. "Automated prediction: Perception, law, and policy." *Communications of the ACM* 55, no. 9 (2012): 33-35.