Centre for Public Impact
A BCG FOUNDATION

ARTIFICIAL INTELLIGENCE
and the future of government

Destination unknown:

Exploring the impact of Artificial Intelligence on Government

Working Paper

September 2017

Acknowledgements

This report was prepared as part of the Artificial Intelligence and Future of Government stream at the Centre for Public Impact (CPI). Joel Tito is the programme manager and principal author of this work and would like to thank Danny Buerkli, Adrian Brown, Farva Kaukab and Nadine Smith from the CPI team for their input, guidance and humour.

The CPI would also like to express our gratitude to the numerous interns, consultants and students who contributed to this project: Anne-Marie Sim, Amalle Wisbech-Heegard, Carla Hustedt, Katie Rose, Kerrie Holmes, Zak Rosentzveig, Adira Levine, Amena Shrafeddin, David Atehortua, George Brighten, Jason Sutjiawan, Rodrigo Mujica-Varas, Olivier Kopff, Jules Wurlod, Simon Mueller and Valerie Ludwig.

The report would not have been possible without the input and reflections from our friends Mike Bracken, Anne Washington and various partners at The Boston Consulting Group who were incredibly generous with their time and expertise.

Finally, thank you to Charles Unwin and Matt Mercer for the often thankless task of making our footnotes make some sense and our words read real good.

Table of Contents

ACRONYMS AND ABBREVIATIONS

AI Artificial Intelligence
APPD Adult Probation and Parole Department
Apps Applications
BSP Bulk Synchronous Parallel
CCTV Closed Circuit Television
CIO Chief Information Officer
COSMOS Collaborative Online Social Media Observatory
CPD Chicago Police Department
CT Computerised Tomography
CV Computer Vision
DARPA Defense Advanced Research Projects Agency
DHS Department of Homeland Security
DoIT Department of Innovation and Technology
DVA Department of Veterans Affairs
EPD Lab Event and Pattern Detection Laboratory
FACE Facial Analysis Comparison and Evaluation Unit
FBI Federal Bureau of Investigation
GLA Greater London Authority
ICP Information Computer Technologies
IT Information Technology
LA Los Angeles
LAPD Los Angeles Police Department
MIT Massachusetts Institute of Technology
MLA Machine Learning Algorithm
MODA Mayor's Office of Data Analytics
MRI Magnetic Resonance Imaging
NHS National Health Service
NLP Natural Language Processing
NSA National Security Agency
NSL National Security Letter
NYC New York City
NYPD New York Police Department
OSTP Office of Science and Technology Policy
PHT Portsmouth Hospitals NHS Trust
PredPol Predictive Policing
PSA Public Safety Assessment-Court
PTSD Post-Traumatic Stress Disorder
RACR Real-Time Analysis and Critical Response
SQL Structured Query Language
TACIDS Tactical Identification System
TfL Transport for London
UC University of California
UCLA University of California, Los Angeles
UK United Kingdom
US United States
VfM Value for Money

I. EXECUTIVE SUMMARY

Artificial intelligence (AI). More than just a phrase or concept, AI is rapidly moving from theory to reality – and this is something we *all* need to be ready for, including governments. But when and how these changes occur, and whether governments can take advantage of AI's many benefits whilst meeting the challenges it brings, largely depends on how policymakers act now.

This paper is intended to help government officials navigate the unfamiliar terrain of this new set of technologies. It is broken down into six sections. Section One **defines and introduces** AI and machine learning algorithms (MLAs). Section Two articulates **how AI can be deployed to help existing government functions**. Section Three examines **how AI can be effective in changing policymaking processes.** Section Four outlines the **challenges for governments** that may prevent them from capturing the benefits of AI. Section Five describes the **risk for governments** of failing to act and of getting things wrong. Section Six proposes a set of **recommendations** to place governments in a position to capture the benefits of AI.

What is AI?

Broadly defined, AI is software that enhances and automates the knowledge-based work conducted by humans. Drawing upon fields as disparate as computer science, cognitive psychology, philosophy, neurology, and others, it is a rapidly expanding field of technological research that is now taking permanent root well beyond the science lab.

How can it help governments?

We address four capabilities of AI that can be deployed to improve both the outcomes governments seek to achieve and the way in which they make policy. These capabilities are:

1. Predictive analytics
2. Detection
3. Computer vision
4. Natural language processing

Predictive analytics have two main goals. The first is to guess the values of an outcome based on the relationship between the outcome and predictors found in a group of observations for which we have complete data. The second goal is to explain predictions of both known and unknown observations.

Detection is used to identify individual data records or patterns (such as a relationship between a mode of hospital care and a health outcome) within massive and complex datasets to identify those that are 'abnormal' or anomalous, thereby providing the user with unparalleled situational awareness.

Computer vision enables the collection, processing and analysis of any information obtained from digital images from various sources (e.g. satellite images, aerial, medical pictures etc.) as well as digital video.

Natural language processing makes it possible for machines to process and understand audio and text data to automate tasks like translation, interactive dialogue and sentiment analysis.

These four capabilities can be applied to improve existing outcomes in health, justice and policymaking and this paper outlines the way in which this has already been done.

What are the challenges?

We investigate the potential challenges to the broader uptake of AI in government. These concern the legal, technical and human resource challenges of deploying AI in government.

The profusion of data silos is an example of the insular management of data systems within government and an important obstacle facing the uptake of AI. In order to build an environment to exploit AI there are a number of key preconditions that must be met. The first enabler is the legal setting. It is required that the law is clear and incentivises data sharing between stakeholders. Secondly, each user (i.e. individual agency) is required to meet a basic technical capacity. Though the design of an e-infrastructure is based on the capability of each agency involved, it is crucial that all parties involved provide strong IT capabilities such that the evolution of that infrastructure can progress efficiently and swiftly. Thirdly, it is necessary that there is sufficient human capacity for adoption to occur, vis-à-vis the talent and human capabilities, as well as the culture, required to embrace the new technology.

What risks does AI present?

We see three broad risks going forward. The first is the risk to government legitimacy should governments continue to fall behind other organisations in society in their use of AI. AI is increasingly being utilised in the private sector and producing improved outcomes. It is clear that governments trail behind in capturing the benefits of AI and in the first part of this section we examine what would happen if this remains the case. We argue that the risk is not just that citizen expectations will continue to outstrip government capacity, but that the very basis for the social contract that gives government its authority to rule may come into question.

The second risk we examine derives from the misuse of AI by government. AI is increasingly being used in decision-making to drive improvements in efficiency and effectiveness. But AI can also systematise inequity by hard-coding bias into decision-making processes. Over-reliance on AI can produce a lock-in to expensive systems, with spiralling costs for government. And in order to fully make use of AI's capabilities, governments may look to centralise their handling of citizens' data, potentially over-extending their powers. As AI increases in complexity, the processes by which AI makes decisions will become increasingly obscure, reducing citizens' power to raise objections.

And in the final part of this section, we examine whether AI will fundamentally change the way government looks, thinks and acts. We suggest that AI is a potentially transformative technology, and we posit that governments need to recognise the ways in which AI could start transforming government if they are to maintain their role in the longer-term.

Recommendations

Our paper concludes by making three broad types of recommendations to government. The objective of these recommendations is to put them in the best position to capture the benefits and overcome the challenges of AI. The three broad recommendations are:

1. **Define needs:** Best practices for identifying departmental need
2. **Build capacity:** Human and technical building blocks required for the uptake of AI
3. **Adapt structures**: Adaptations required to existing cultural, regulatory and legislative environments.

How can this paper be used?

Government officials can benefit from this paper in three ways. First, it aims to provide a concise account of AI and machine learning algorithms (MLAs) that can be understood by a newcomer. The field of data science is wide, technical, and advancing rapidly into new applications. Terminology is not standardised and it can be difficult for a non-expert to understand how to use

MLAs in their field of expertise. To that end, the paper provides a clear overview of AI's integration into public policymaking and service delivery.

Second, it aims to provide an assessment of the nature and value of the opportunities for the uptake of AI in government so as to identify current needs that could be met using this technology. It also articulates the risks associated with failing to capture these benefits. Lastly, the paper seeks to provide a set of actionable recommendations that governments can follow in order to take advantage of this transformative technology.

There is a clear demand for this technology. Government operating models have remained in near-stasis for decades, but many have begun to make the transition to e-governance in order to make public policy more evidence-based, dynamic and valuable. While there is still significant progress to be made, there are clear indications that governments have started to take some preliminary steps towards digital government: these current trends have helped lay the foundation for the use of AI in the public sector.

II. INTRODUCTION TO ARTIFICIAL INTELLIGENCE

A. What is AI?

Drawing upon fields as disparate as computer science, cognitive psychology, philosophy, neurology, and others, AI is a rapidly expanding field of technological research.[1] Given its overlaps with advanced statistical analysis and its evolving capabilities, definitions of AI vary. Russell and Norvig's seminal AI textbook, *Artificial Intelligence: A Modern Approach,* defines AI as technology that can think humanly, act humanly, think rationally, or act rationally.[2] Other definitions, which fall into these categories, include: "The study of how to make computers do things at which, at the moment, people are better"[3] and "The study of mental faculties through the use of computational models".[4] An AI textbook by Poole and Mackworth defines AI as "the field that studies the synthesis and analysis of computational agents that act intelligently," where intelligently refers to acting appropriately, exhibiting flexibility, and learning from experience.[5]

Broadly defined, AI is software that enhances and automates the knowledge-based work done by humans.[6] This software is powered by machine learning. Depending on the type of work it is being applied to, AI can be described as narrow or general.[7] Narrow AI describes the application of AI to individual tasks that are repetitive or based on patterns, whilst general AI aims to create a machine that is capable of performing all of the intellectual tasks that a human brain can.[8]

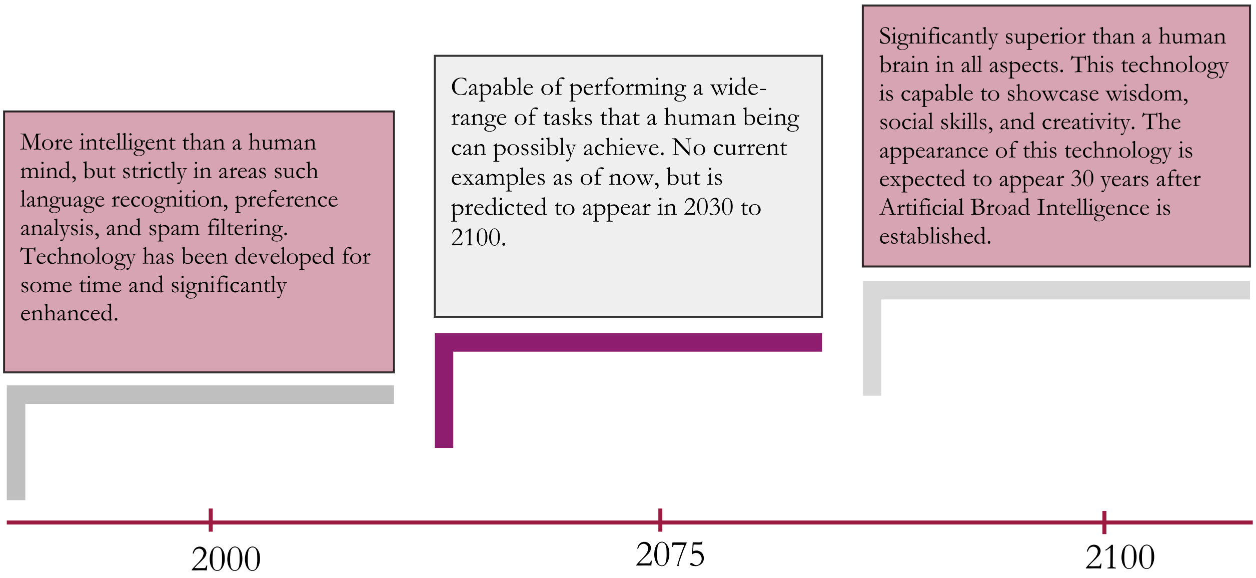
Artificial general intelligence will reason, learn, and problem-solve in complex and changing environments as well as humans do.[9] These abilities do not yet exist, and are not estimated to emerge before 2030.[10] Owing to the difficulties in achieving general AI, some estimates predict that these capabilities will not develop for over a century.[11]

---

1 Russell, S. and Norvig, P. (2013)
2 *Ibid.*
3 Rich, E. and Knight, K. (1991)
4 Charniak, E. and McDermott, D. (1985)
5 Poole, D. and Mackworth, A. (2010)
6 Griffin et al. (2016)
7 This is also referred to as weak or strong AI
8 *Ibid.*
9 Urban, T. (2015)
10 Holdren, J., et al. (2016); Urban, T. (2015)
11 Holdren, J., et al. (2016)

Following the development of artificial general intelligence, some analysts forecast the emergence of artificial super-intelligence, a type of AI that far surpasses human intellect and abilities in nearly all areas.[12]

Figure I: Predicted Artificial Intelligence Development



Source: EIU (2016)

Today's AI falls into the narrow AI category, and this is the focus of this paper. By concentrating on narrow AI, the paper avoids some of the more speculative arguments about AI's transformative power. Yet, even without veering into the domain of general AI, the impact on governments of AI adoption will be enormous. Its uses have become increasingly widespread and its reach pervasive.[13] AI capabilities include reasoning, machine learning, robotics, natural language processing, object perception, information storage and retrieval, and speech and handwriting recognition. These tools help power self-driving cars, automated translation, search engine results, and game-playing robots, among many other applications.[14]

One key type of AI is machine learning, a process in which computer programs analyse data and use the newfound knowledge to inform a decision or prediction.[15] This is explored in more detail in the following section.

12 Holdren, J., et al. (2016); Urban (2015)
13 Urban, T. (2015)
14 Russell, S. and Norvig, P. (2014)
15 Copeland, M. (2016)

i. What is a machine learning algorithm?

The field of data science is wide, complex and continuously advancing into new applications. As a result, terminology is not standardised and definitions reflect common usage rather than accepted definitions.[16] Many of the terms used to describe the field – big data analytics, MLAs, AI, data mining, cognitive computing – overlap in terms of the problems that they solve and the techniques that they use to solve them.[17]

That said, it is useful to conceptualise the field in some way. Loosely conceived, MLAs are tools for solving learning problems. These problems require the agent, human or machine, to improve their performance (measured, for example, by accuracy) in executing a given task (e.g. detection of fraud) through exposure to a set of historical examples (e.g. credit card transactions).[18] Before the advent of computer technology, human brains were the primary tool used for solving these problems. Now, MLAs are able to do so.

There are two broad categories of learning problems that MLAs solve: supervised and unsupervised.

Supervised learning algorithms

In supervised learning, MLAs are used to find the most accurate way (known as a function) to predict an outcome (y*) based on previous examples of relationships between inputs (x*) and that outcome. The user tells the MLA which outcome to predict and the learning experience is a labelled[19] historical dataset, which are group observations (x,y pairs) that have ground-truth.[20]

An example of a very simple historical dataset is one that contains 2,000 individual properties with data on the outcome that the user is interested in, such as whether or not a convicted individual reoffends within two years of being released from prison, and other characteristics, such as age and gender. The objective of this is to find the function that makes the most accurate prediction of the outcome in both the data that is used to find the function (known as the sample or training dataset) and, more importantly, any new and applicable dataset.

---

16 Varian, H. (2014)
17 Neill (2012)
18 Jordan, M. and Mitchell, P. (2015)
19 Labelled data is usually structured (in rows and columns) and, as the name suggests, has labels for the data so as to "tag" each piece of information. Unlabelled data is defined as such because of the lack of "explanation" to categorise the data. It is usually unstructured (e.g. images and videos).
20 Jordan, M. and Mitchell, P. (2015)

For instance, age and gender may be two variables that are strongly associated with recidivism individually, but that are more strongly associated with recidivism when acting together.[21] The MLA, then, might first find a strong association between recidivism and individuals who are males as well as a strong association between recidivism and individuals in their late teens, but after being exposed to enough training instances it might find an even stronger relationship when an individual is male *and* in his late teens. In this way, the function gradually improves itself as it is exposed to new data and finds different functional forms that improve its accuracy.[22]

If the user wants to predict a nominal attribute (something measured as a number of discrete categories), such as the condition that a patient's symptoms correspond to a certain condition, then a classification algorithm is most appropriate. If the attribute is numeric (measured on a continuous scale), such as the likelihood of a patient's condition deteriorating into something more severe, then the user should run a regression algorithm.[23] It is important to note that a function establishes only correlation – not causality. The function does not allow us to infer the causal impact of explicitly assigning a change in one attribute, the number of police in a precinct, for example, on the attribute we are trying to predict. All the function allows us to do is to infer that there is an association between the two.

Unsupervised learning algorithms

In contrast to supervised MLAs, unsupervised algorithms work with data that is unlabelled. The algorithm learns on its own: the user does not tell the algorithm what to look for. Instead, the MLA is programmed to do things like discover the structure of the whole dataset (a task known as modelling), or identify interesting subsets of the dataset, such as anomalous records or relationships (a task known as detection).[24] An example of modelling might be the identification of different groups of prisoners within a dataset on convicted felons or the extraction of different features in unlabelled image data, while an example of detection might be the identification of

[21] Footnote to be updated: Labi, Nadya (2012); Berk., R. and Hyatt, J. (2015)

[22] The goal of supervised learning is to find a function that gets good out-of-sample predictions. If the function is accurate for all of the observations in the sample dataset but is inaccurate on observations in new datasets, then the function suffers from overfitting. Conversely, when the function only considers simple relationships between the attributes and the outcome, the function is considered to be underfitted. The MLA user can avoid these problems by using techniques, such as out-of-sample cross validation and regularisation, that penalise the out-of-sample variance – or the "wrongness" of the MLA outside of the training dataset – of the function.

[23] Neill, D. (2013)

[24] *ibid*

fraudulent transactions or anomalous patterns of care and health outcomes. Examples of unsupervised MLAs include clustering and subset scanning algorithms.

III. How can AI help existing government functions?

According to Martin Painter's analysis of public sector reform, the last fifteen years has seen a "paradigmatic realignment of state, markets and civil society" which has been galvanised by open data initiatives, self-governance, social enterprise, decentralisation and network governance.[25] Data-driven techniques have also contributed to this shift. The current epoch has been interchangeably defined as the 'information age', 'digital age' or 'network society'. This era is based on an information economy and the evolution in data management has put governments under pressure to adopt e-government strategies. From the incorporation of information computer technologies (ICTs) to the digitalisation of some services, the public sector has sought to modernise and improve the way decisions are made and services delivered in order to keep pace with private sector developments.

This trend has seen governments attempt to transition from a programmatic model of service delivery to a citizen-focused model.[26] Alongside this, service delivery has had to advance from an approach focused solely on service quality to a model that emphasises the delivery of better outcomes more efficiently.[27] As will be shown in the following section of this paper, the benefits generated from AI's sophisticated data analytics in government are vast. Current applications have tended to be focused on service delivery but the use of AI to inform decision-making across government will provide an important opportunity for public institutions to achieve value-for-money and broader social benefits. With many governments making the transition to e-government, public bodies are quickly realising that technology can facilitate more efficient and effective service delivery, dramatically improving public sector value. AI capabilities clearly outstrip human capabilities in certain areas, such as data processing ability, and improvement in service delivery outcomes has been demonstrated: AI has been shown to speed up services, improve on human accuracy, reduce the number of people necessary to fulfil specific tasks and organise sophisticated ideas via expertise analysis.

There have been many attempts to put a number on the value of big data and machine learning to governments. An estimate from the Policy Exchange, a UK thinktank, estimates that £16-33 billion (around $24-48 billion) can be saved by the UK government alone by using these technologies, which amounts to £250–£500 (around $365-735) per capita gains. The Centre for

[25] Painter, M. (2010). Footnote to be updated.
[26] Chang, A. and Kannan, P. (2008)
[27] Daly, E. and Singham, S. (2012)

Economic and Business Research estimates the value over the period 2012-17 of using big data in government to be £20.4 billion (around $30 billion).[28]

While estimates of the monetary value of the application of AI to government can be instructive, they ignore the social value that big data and AI can create. Social values are difficult to assess in monetary terms but are relevant for generating a comprehensive assessment of the value of AI. One example of the social value that AI can generate is in the potential for reducing the inconsistency and bias present in human decision-making.

For example, within the justice system, Danziger et al illustrate inconsistency through a study of judicial rulings around food breaks.[29] They find that the percentage of favourable rulings drops gradually from approximately 65 percent to nearly zero as a judge approaches a food break, returning abruptly to approximately 65 percent after a food break. Kahneman calls this the 'noisiness' of human decision-making;[30] it is inherently randomly inconsistent. Biases – both cognitive and social – also infiltrate human decision-making. Cognitive biases, a concept introduced by Tversky and Kahneman, are the result of systematic errors associated with heuristic rules.[31] 'Anchoring bias' is one example, wherein humans rely disproportionately on the first piece of information they encounter, rather than weighing all information dispassionately. Social biases are based on prior beliefs and worldviews, and sometimes manifest in discrimination. Algorithms have the potential to eliminate both noise and bias from decision-making.[32] For example, algorithms can weigh all inputs exactly as instructed, helping to avoid anchoring bias.

This section is broken down into two parts. The first part articulates four capabilities of AI that can enhance existing activities. The second part looks at how these capabilities could improve existing government outcomes, with a specific focus on health and justice.

A. Four capabilities that will enhance government's existing activities

Whilst the capabilities that AI presents are wide-ranging and interconnected, here we present four

[28] Yiu, C. (2012). Footnote to be updated.
[29] Danziger, S. et al (2011)
[30] Kahneman, D. (2016)
[31] Tversky, A. and Kahneman, D. (1974)
[32] Kahneman, D. et al (2016)

specific AI capabilities that can be deployed to improve both the outcomes governments seek to achieve and the way in which they make policy. These capabilities are:

1. Predictive analytics
2. Detection
3. Computer vision
4. Natural language processing.

These capabilities are neither mutually exclusive nor exhaustive, and as the following section shows, AI solutions may exploit multiple capabilities in tandem for optimum effectiveness. Nevertheless, we believe that this delineation of capabilities is a useful means of illuminating the specific features and associated benefits that AI technologies present. Before demonstrating how these capabilities can be applied, it is important to first define what they are.

i. Predictive analytics

Much of the research that policymakers use focuses on causality. However, there are policy questions that do not necessarily require knowledge about causality; correlation between independent data points can often be sufficient to inform appropriate interventions. In such cases, better predictions can enable governments to implement preventative and/or personalised policies.[33]

By using massive training datasets to select attributes that best predict an unknown function such as an inmate's risk of reoffending, AI can help both policymakers and frontline civil servants to make predictions in a way that is more comprehensive and less subject to human bias. There are two main goals of prediction. The first is to guess the values of an outcome based on the relationship between the outcome and predictors found in a group of observations for which we have complete data.[34] For instance, a clinician might want to predict the length of a patient's stay based on his or her health profile. The second goal is to provide predictions on the basis of incomplete information. For instance, predictive analytics could be used to map a complex decision tree of all possible outcomes, which might then provide a more manageable basis for human diagnosis.

[33] Kleinberg, J. et al (2015)
[34] Neill, D. (2012)

ii. Detection

Another important set of tasks that AI can assist with is detection tasks. In these tasks the goal is to automatically detect individual data records (such as a fraudulent transaction) or patterns (such as a relationship between a mode of hospital care and a health outcome) within massive and complex datasets to identify those that are 'abnormal' or anomalous,[35] thereby providing the user with unparalleled situational awareness. Before the user runs the algorithm, he or she does not know which subset of the data will be returned as anomalous. By directing the user's attention to a subset of the data – and, in the case of anomalous pattern detection, characterising what is interesting about that subset – this process can be used to solve these problems and inform prescient action.

There are two main benefits to using AI for these detection tasks. First, detection MLAs help government officials to identify important patterns that might be causal. We particularly care about this in the medical domain: for example, a hospital may want to identify anomalous patterns between modes of care and health outcomes that can then be verified (using experimental research) and acted upon in the future. Second, detection MLAs help government achieve an unprecedented level of situational awareness. It isn't always imperative that governments understand what is causing an abnormal event, only that that they know the event has occurred. According to Brett Goldstein, ex-CIO of the City of Chicago, the ability to answer the question, "Am I in a normal state?" is something that is "under-realised but enormously valuable".[36] For example, a city official may want to know whether, managing for the variables that affect a system, there is an abnormality in a specific area of the city.

iii. Computer vision

Computer vision (CV) enables the collection, processing and analysis of any information obtained from digital images from various sources (e.g. satellite images, aerial, medical pictures etc.) as well as digital video. In such contexts, the goal of an AI is to find unsupervised methods of feature recognition, identification of objects, actions or characteristics, describe content and, overall, automate labour-intensive cognitive tasks that would usually require human supervision. Traditionally, the public sector has used CV for things like traffic control (e.g. automated number plate identification technology) and policing (e.g. fingerprint matching). However, more advanced

[35] Neill, D. (2013)
[36] Goldstein, B. (2015)

applications like the analysis of MRIs or CT scans has been limited due to high inaccuracy rates and low processing speed. Deep Neural Networks have seen the accuracy of CV leapfrog human level ability in certain areas.

iv. Natural language processing

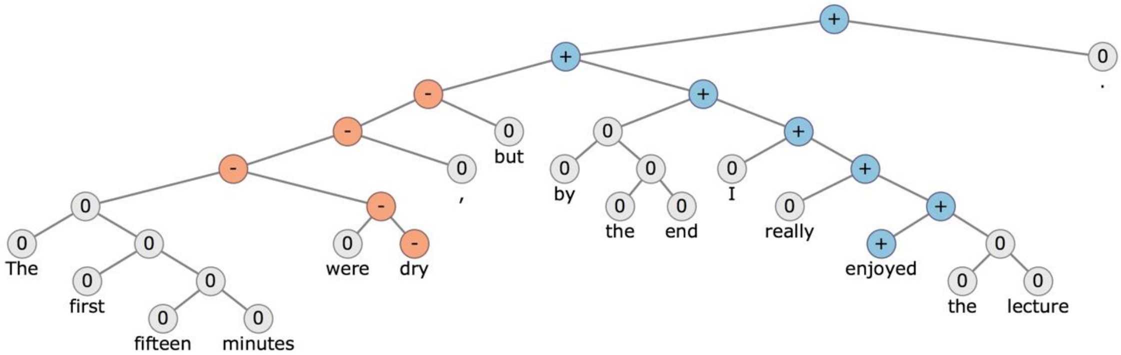
Natural language processing (NLP) makes it possible for machines to process and understand audio and text data to automate tasks like translation, interactive dialogue and sentiment analysis. Machine learning has enabled NLP research to evolve from lexical semantics (meanings of individual words) to compositional semantics (meanings of sentences) and even narrative understanding, which enables machines to process wells of unstructured text data to derive meaningful insights.

Our research indicates that government interest in testing these technologies has concentrated on two areas: first, to mine sentiment and expert content regarding citizen preferences and information related to policy propositions or implementations; secondly, highly advanced applications of NLP are being deployed for surveillance and biometric identification.[37]

Sentiment analysis is the use of AI to extract information from unstructured data (such as text data from social media (e.g. social networks, blogs) or audio data from government hotlines on citizen preferences about existing or proposed policies at the local, regional or national government levels). For example, the diagram below depicts analysis of the sentence: "The first fifteen minutes were dry, but by the end I really enjoyed the lecture". Red nodes indicate that the model assigned a negative sentiment to the underlying word or phrase, such as in the case of "were dry". The phrase "by the end" is correctly assigned a positive sentiment – indicated by the blue node.

Figure II: Decision Tree for Sentiment Analysis of Compositional Semantics

[37] Rahmes, M. et al (2013)



Source: Abate (2013)

B. Applications to existing government functions

We now turn to the way in which the four capabilities discussed above can be deployed to enhance the quality of existing government outcomes. To narrow our focus, we have selected two fields to apply these capabilities to: justice and health. We will deal with each in turn.

i. Justice

In order to examine the impact of AI on justice, it is important to see it as a system rather than restrict our analysis to one element. Accordingly, we describe below the application of the above capabilities to the **police**, **courts** and **corrections**.

Police

One of the primary responsibilities of the police is to keep citizens safe. Many police departments have traditionally relied on information-constrained, ad-hoc and subjective assessments by police officials to direct their activities. With AI, police departments have the ability to use the predictability of criminal activity to their advantage. Over the last decade, a number of large local police departments in the US, including the Los Angeles Police Department (LAPD), the Chicago Police Department (CPD) and the New York Police Department (NYPD), have developed an MLA-driven approach to policing known as 'predictive policing'. This technology has primarily been used to predict crime hotspots but departments have also developed applications for the prediction of gun violence and adverse police interactions.

In 2011, the Santa Cruz Police Department piloted a predictive analytics tool, PredPol, that had been developed by researchers at UCLA, Santa Clara University and UC Irvine specifically for law enforcement.[38] This tool uses data on location, place and time of individual crimes to predict hotspots of crime, which are areas of 500 square feet in which a crime is probable within the next twelve hours. LAPD adopted PredPol in selected neighbourhoods in 2011 and in that year, burglaries fell by 27 percent. In 2012, property theft fell by 19 percent and, relative to LA neighbourhoods in which PredPol had not been introduced, the overall crime rate fell by 13 percent.[39] On the back of this success, LAPD has widened its use of the tool and by 2014, one in three of the LAPD's geographic policing divisions was using PredPol to identify crime hotspots.[40] These trends provide preliminary evidence that predictive policing has reduced crime rates and, following a recent report by the LAPD, driven efficiency gains. By predicting more than twice as many crimes as experienced analysts (4.7 percent of crimes compared to 2.1 percent), PredPol has saved the LAPD $9m per year.[41] The use of PredPol has also produced positive feedback in the way the department collects and uses its data. Realising the power of including different data sources in the model, LAPD's Real-Time Analysis and Critical Response Division (RACR) has begun to use historical and real-time data from closed-circuit TV cameras, gunshot detectors and license-plate readers to augment the prediction of crime hotspots,[42] These data sources are unstructured and require more complex algorithms (often unsupervised learning algorithms) for processing before they are used in a predictive model.

Since 2009, the CPD has worked with Carnegie Mellon University to develop CrimeScan, a tool that, like PredPol, predicts future crimes. Unlike PredPol, CrimeScan's approach involves two stages. The first stage involves the detection of emerging geographic clusters of leading indicators such as minor crimes, 911 calls or 311 calls for service. These clusters of instances, identified by indicator values that exceed expectations, are then used to predict violent crime based on the assumption that areas which are closer to a significant cluster of any of the monitored leading indicator are assumed more likely to have a spike in violent crime within the next week.[43] The prediction algorithm used by CrimeScan then learns which combinations of data sources and variables produce the most accurate predictions of violent crime. This allows police departments to identify hotspots of violence as they emerge, which leads to greater prediction accuracy. Predictive policing has enabled police departments across the US to better anticipate crime and deploy more targeted police interventions. In this way, law enforcement can adopt a preventative

38 Goldsmith, S. (2014)
39 *Ibid.*
40 Berg, N. (2014)
41 Mohler, G. et al (2015)
42 DeSouza, K. and Jacob, B. (2014)
43 Neill, D. (2012)

rather than a reactive approach to solving crime, which improves the efficiency and efficacy of their service.

Integrated with Closed Circuit Television (CCTV), CV software systems are capable of counting, measuring speed, and monitoring direction. These functions are particularly useful for policing large human gatherings like concerts, sporting events, or central transport stations. This technology can recognise different behaviours and trigger an alarm for police or even intelligence departments to act accordingly on a user-defined rule, such as "vehicle parked in unauthorised zone" or "unattended bag on platform" and even "crime suspect identified".

CV is getting so precise, and the processing power of the hardware that supports it so great, that even highly complex operations like facial recognition can be automated currently with human level accuracy.[44]

The FBI created the Facial Analysis Comparison and Evaluation Unit (FACE), which provides investigative support using face recognition software to compare facial images contained within government systems (e.g. the Department of Motor Vehicles' photographs) with pictures of missing persons and fugitives.[45] The success of the pilot programme led to its expansion to eleven states, and the creation of new applications such as the "Tactical Identification System" (TACIDS) in San Diego. TACIDS provides law enforcement officers with smartphones that can run photos taken in the field against the sheriff's mugshot database. In the UK, Leicestershire police started a trial in April 2015 of a similar application called NeoFace, which has been proved to find the real identity of pictured suspects in 45 percent of cases and within seconds. NeoFace has already been used on over 200 occasions during trials and is used on a daily basis in police investigations, resulting in efficiency savings by cutting out labour-intensive and lengthy searches of criminal mugshot databases.[46]

Courts

Risk assessments are used across the court system but have traditionally been made in an ad-hoc, information-constrained and subjective way. While the convicted individual's crime and record are useful pieces of information in an assessment of his or her 'future dangerousness', a risk assessment based on this information may overlook a number of important predictors. Like their

[44] Krizhevsky, A. (2012)
[45] Klontz, J. (2013)
[46] Guo, G. and Wang, X. (2012)

implementation in police departments, MLAs promise to not only improve the efficiency of the courts – namely by increasing the speed at which bail decisions are made – but also their efficacy.

A selection of jurisdictions in the US is trying to automate and debias risk assessments by using MLAs across the court system. Some systems already use risk assessment algorithms to identify which prisoners to release on parole and set bail. Indeed, the decision whether to detain or release arrestees as they await adjudication of their case is a decision that depends on a prediction about the arrestee's probability of committing a crime during this period or likelihood of skipping court. For instance, the Laura and John Arnold Foundation commissioned the development of Public Safety Assessment-Court (PSA), an algorithm that uses data on an individual's age, his or her criminal record, and previous failures to appear in court to assess the likelihood that an individual will commit a crime or fail to reappear in court. This is currently being used by states like Arizona and New Jersey and cities like Chicago and Pittsburgh and has seen some initial success: according to the Foundation, lower crime rates and jail populations have been associated with the introduction of the tool.[47] Researchers at Harvard and the University of Pennsylvania have also developed a similar tool.[48]

Pennsylvania is taking steps to apply algorithmic risk assessment to sentencing.[49] A state commission is drafting a plan that would allow judges to use an MLA-based risk assessment tool to assist with sentencing decisions. Berk and Hyatt, two researchers from the University of Pennsylvania, have recently published an application that provides some insight on how Pennsylvania might use MLAs to inform sentencing decisions. In this application, the authors argue that MLAs can provide an accurate and cost-adjusted prediction of the risk that an individual poses to society. More specifically, Berk and Hyatt propose a supervised prediction algorithm that performs better than conventional modelling approaches and takes into account the asymmetric costs of forecasting errors: in this application a false negative (predicting an individual as low risk when he is actually high risk) is much more costly to society than a false positive (predicting an individual as high risk when he is actually low risk).

Corrections

Recidivism is a malfunction of many criminal justice systems. In the UK, the reoffending rate has

[47] Dewan, S. (2015)
[48] Kleinberg, J. et al (2015); Berk, M. and Hyatt, J. (2015)
[49] Barry-Jester, A. et al (2015).

fluctuated between 26 percent and 29 percent since 2003.[50] In the US between 2005 to 2010, around two-thirds of released prisoners were rearrested within three years.[51] The response of governments to recidivism is to place at-risk individuals under community correctional supervision. However, parole and probation departments must balance fiscal prudence with public safety.

Fortunately, recidivism is not always a random variable and can be predicted using AI. Pennsylvania has a record on recidivism that is representative of the national average: one in three inmates is arrested again or reincarcerated within a year of being released (PR Newswire, 2015). In 2006, Philadelphia's Adult Probation and Parole Department (APPD) partnered with the University of Pennsylvania to develop a prediction algorithm to forecast the risk of recidivism for individual probationers.[52] This tool has been used in Philadelphia for seven years and has achieved an average accuracy rate of 66 percent across all probationers. Similar models are being developed at MIT, which suggests that this approach to recidivism management may scale.[53]

ii. Health

The use of AI for prediction in public healthcare has primarily focused on patient-level clinical support decisions. In particular, MLAs have been used to help with patient diagnosis, treatment and monitoring. While many of these patient-level applications are still at the trial stage, it is conceivable that AI could also be applied to predictions at the hospital level.

The application of AI to patient-level predictions, commonly grouped together under the term 'personalised medicine', involve the use of MLAs to improve clinician decisions on the diagnosis, treatment and monitoring of their patients. This application is a rapidly growing area of healthcare: between 2006 and 2014, the number of personalised medicine diagnostics and treatments available increased from 13 to 113 worldwide.[54] The NHS is a public healthcare system that is increasingly supporting the use of such algorithms to personalise and subsequently improve healthcare service delivery, not least through the personalised medicine strategy announced by NHS England's National Medical Director in his Five Year Forward View in September 2015.[55]

---

[50] Ministry of Justice (2015)
[51] National Institute of Justice (2014)
[52] Ritter, N. (2013)
[53] Zeng, J. et al. (2016)
[54] PMC (2014)
[55] NHS England (2015)

Patient treatment

Patient treatment is a prediction problem that doctors face every day. The efficacy of a treatment for a given condition differs across patients because of genetic and phenotypic variation. Indeed, many currently available drug treatments are only effective for 30-60 percent of treated individuals. Traditionally, doctors have made subjective assessments of the best treatment regimen for a given patient, but with AI doctors are able to segment patients into population subgroups according to their genotype and phenotype (known as molecular signature matching) and tailor the treatment type and dosage to maximise efficacy and minimise side effects.[56] This approach enables clinicians to make comparisons of treatment options that were too complex and time-consuming to be made manually before.

One of the first applications of MLAs for personalised treatment was developed by the MD Anderson Cancer Center at the University of Texas in partnership with IBM Watson.[57] This application, called the Oncology Expert Advisor, uses massive amounts of patient data and medical literature to provide oncologists with evidence-based care decisions on first-line therapy that are tailored to each patient. This application has proven to be a successful way of quickly assessing the best treatments for an individual patient based on the latest evidence: the overall accuracy of standard of care recommendations in 200 test leukemia cases was over 80 percent.[58]

While personalised treatment is still very much in the trial stage, it has significant potential because of the vast quantities of patient genomic data that is coming online through initiatives such as the 100,000 Genomes project. Sir Malcolm Grant has claimed that such advances in the data capabilities of the NHS could allow "clinicians to make more effective use of expensive drugs, such as those used in chemotherapy, by attuning them to tumour DNA and then monitoring their effect through a course of treatment".[59] Indeed, researchers at the Precision Cancer Medicine and Centre for Molecular Medicine at the University of Oxford are using a £132m investment to develop more precise and personalised diagnostic and treatment tools.[60] The application of AI to personalised treatment has been shown to have value in the UK context. Working with publicly available NHS prescriptions data, Open Health Care UK analysed prescribing patterns in the NHS

[56] Manyika, J. (2011)
[57] Stein, A. (2014)
[58] *ibid.*
[59] Griffin, C. et al (2016). Footnote to be updated.
[60] University of Oxford (2014)

and found an average of £27m each month of 2012 was unnecessary expenditure on statins, a drug used to prevent cardiovascular problems.[61]

CV is helping doctors make diagnoses and spot abnormalities in medical images. Enlitic, a deep learning healthcare company, can analyse medical images such as X-rays, MRIs, or CT scans for trends in the data or anomalies in individual images. The organisation trained their AI with unstructured lung CT scans data, aiming to diagnose potentially cancerous growths.[62] The results were notable: diagnoses made by four of the world's top human radiologists had a false negative rate (missing a cancer diagnosis) of 7 percent, whereas Enlitic's MLA produced zero false negatives. Human radiologists had a false positive rate (incorrectly diagnosing cancer) of 66 percent, while Enlitic had a false positive rate of 47 percent.[63] The accuracy of this machine has illuminated tremendous potential for growth in this field and the capacity for MLAs to radically improve the way medical diagnosis is carried out in public health systems.

In late 2014, the US Department of Veterans Affairs (DVA) began to explore how AI could help veterans' affairs doctors rapidly sift through electronic medical records for treatment and research data that could support clinical decisions. In particular, the DVA began a pilot of IBM's Watson system to investigate the value of a comprehensive decision support system for the care of veterans who suffer from PTSD. The outcome of this application is drastically increased diagnosis accuracy. Given that one in five diagnoses are incorrect in the US every year and nearly 1.5 million medication errors are made in the US per annum,[64] the importance of a tool that can facilitate decisions is profound, both in terms of cost savings and improved treatment.

Watson Health supports medical professionals as they make decisions. For example, a physician can use Watson to assist in diagnosing and treating patients. To begin with, a doctor might pose a query to the system, describing symptoms and other related factors. Watson begins by parsing the input to identify the critical pieces of information. It then mines the patient data to find relevant facts about family history, current medications and other existing conditions. It combines this information with current findings from tests and instruments and then examines all available data sources to form hypotheses and test them.[65]

---

[61] Thwaites, E. (2012)
[62] Dorrier, J. (2015)
[63] *ibid.*
[64] IBM (2017)
[65] IBM Watson (2014)

Watson can incorporate treatment guidelines, electronic medical record data, doctors' and nurses' notes, research, clinical studies, journal articles and patient information into the data available for analysis. It will then provide a list of potential diagnoses along with a score that indicates the level of confidence for each hypothesis.

Patient monitoring

While in hospital, some inpatients experience a deterioration in their medical condition or suffer from a complication of their illness such as a life-threatening blood clot. To protect patients against these risks, clinicians have traditionally made subjective assessments of the probability of individual patients suffering a deterioration in their condition based on ad-hoc interpretations of vital signs such as temperature and blood pressure. By applying MLAs to this prediction problem in the form of Early Warning Systems (EWS), clinicians are able to automate patient risk assessments and, in doing so, improve their accuracy and comprehensiveness which, in turn, will prevent the number of deaths from deterioration. Based on the academic literature, the main set of algorithms that EWS use are a set of time-series classification algorithms.[66]

A number of NHS Trust Hospitals have implemented this application. Funded with £1.8m by the Wellcome Trust, a foundation, and the UK Department of Health, the University of Oxford Trust has developed Hospital Alerting via Electronic Noticeboard (HAVEN), which is a hospital-wide EWS that uses prediction MLAs to make continuous risk-assessments of individual patients based on real-time vital signs data and other sources such as patient descriptors (e.g. age, admissions history) and laboratory results.[67] Piloted in Oxford University Hospitals and the Portsmouth NHS Trust, this system provides expert clinicians with real-time access to the risk-assessment scores of every patient in the hospital, which enables them to identify and prioritise at-risk patients and deliver the appropriate treatments. Using funding from the same grant source, the Birmingham NHS Trust has also piloted an EWS called RAPID, the Real-Time Adaptive and Predictive Indicator of Deterioration, that uses a combination of prediction and detection algorithms to alert clinicians to deteriorating patients.[68]

The health sector has also explored the application of facial expression recognition software, a

[66] For example Saria, S. et al (2010).
[67] Watkinson, P. et al. (2015); Health Innovation Challenge Fund (2017)
[68] Birmingham Children's Hospital (2015)

subset of facial recognition, which is already being used in the private sector in consumer behaviour research, usability studies and market research. MLA-trained software can learn to recognise facial expressions, from the most obvious to hidden micro-expressions, with wide-ranging applications from psychological analysis to hospital patient monitoring.[69]

A team of paediatricians from the Institute for Neural Computation, together with Emotient, an emotion measurement company, developed a facial expression recognition CV model for assessment of children's postoperative (laparoscopic appendectomy) pain. The team argued that human pain assessment is subject to bias and often under-recognises pain in youth, whereas facial expressions are a reliable biomarker. The model eliminates human bias: it estimated youth self-reported pain better than nurses. In assessing pain severity, the model was as accurate as parents' assessments for the 50 children that took part in the experiment. And labour intensive patient monitoring could also be partially automated – the model's assessment of pain versus no pain was highly accurate for ongoing and transient pain conditions.[70]

Hospital-level applications: Innovations in healthcare practices

Hospital managers want to know which patterns of care are effective and which are not: for instance, whether the hospital is doing things that increase the risk of the spread of hospital acquired illnesses. Since the hospital manager does not know which patterns to look for, this is a detection problem that can be solved by unsupervised MLAs.

MLAs can help hospitals identify anomalous patterns of care and patient health outcomes within massive quantities of healthcare data that are not manually detectable. Indeed, within a hospital, there is huge natural variation in care practices.[71] Different clinicians will administer different types and dosages of medications for a given condition, while different hospital staff will adhere to different degrees of hand washing, isolation precautions and physician orders. This variation in care practices may have significant but as yet undetected impacts on patient outcomes, such as mortality and morbidity rates or hospital readmissions. For example, a group of patients with coronary heart disease might suffer fewer complications if given drugs X and Y before angioplasty.

By using MLAs to detect these patterns, managers can then make decisions about which care

[69] Li, X. et al (2015)
[70] Murdoch (2013). Footnote to be updated.
[71] Neill, D. (2013)

practices to drop or change due to unusually negative patient outcomes and which care practices to keep or scale up due to unusually positive patient outcomes. In addition, the patterns that detection MLAs find can be treated as hypotheses (i.e. this pattern of care causes a change in this health outcome). These can then be evaluated either by medical expertise or, more rigorously, in (quasi)-experiments for potential use in future healthcare practices outside of the hospital in which they were identified.[72] While nascent, this application has been developed by Baylor College of Medicine in partnership with IBM Watson. In this application, researchers at Baylor used MLAs to generate better tumour suppression hypotheses using the massive (around 70,000 papers large) medical literature on 'p53', a protein that can be targeted to suppress tumours.[73]

Nation-level applications: Disease surveillance

At the national level, the value of detection in public healthcare lies in providing situational awareness. Detection MLAs have the potential to detect emerging disease outbreaks quickly and accurately, which enables a rapid response from healthcare officials. This is an important tool to have: the U.S. Defense Advanced Research Projects Agency (DARPA) estimate that a two-day gain in detection time and public health response could reduce fatalities by a factor of six.[74]

The NHS has tested disease surveillance systems that rely on machine learning methods in a number of hospitals. In collaboration with The Learning Clinic, Portsmouth Hospitals NHS Trust (PHT) built VitalPAC IPC, a permanent EWS of potential infection outbreaks, and applied it to the detection of norovirus outbreaks.[75] Norovirus outbreaks impose significant costs on hospitals: in the UK, they affect 13,000 patients and 3,400 staff every year, which leads to 8,900 days of ward closure and 15,500 lost bed-days. This imposes an annual cost of £41.5m on the NHS.[76]

At PHT, VitalPAC IPC is used on routine and other clinical information to identify patterns of symptoms that might be related to a norovirus outbreak. Once a possible outbreak is detected, the system notifies the Infection Control Team, who can intervene with preventive measures such as increased hygiene measures and quarantine. Outbreaks of norovirus at Queen Alexandra Hospital fell from 20 per year to just five over three years after its introduction – a 95 percent reduction, compared with 15 percent and 28 percent for Wessex and England, respectively. If this system

[72] *ibid*
[73] Spangler, S. et al (2014)
[74] Neill, D. (2012)
[75] Mitchell, C. et al (2015)
[76] *ibid.*

were to be implemented across the NHS, it is estimated that annual savings could be as high as£38.5m.[77] There is huge amount of potential value to be gained by scaling this application up to the national level.

[77] McCall (2015). Footnote to be updated.

IV. HOW CAN AI AFFECT THE POLICYMAKING PROCESS?

In this section we look at the way in which AI can assist policymakers. Multiple frameworks exist to conceptualise the policy process. The most common model is a multi-stage framework of policymaking called the policy cycle.[78] While the policy cycle cannot capture all of the complexities of policymaking, it is a useful model for illuminating some of the contributions AI can make throughout the policymaking process. Its separation of policy stages provides a clear and demarcated framework through which to interrogate AI's various opportunities for contribution.[79]

Howlett and Ramesh's (2003) five-stage framework of policymaking, used in their *Studying Public Policy* textbook,[80] is included below:

Agenda setting: Identifying a problem and bringing it to the attention of policymakers.[81]
Policy formulation: Considering available policy options and crafting solutions to identified problems.[82]
Policy decision-making: Government choice of action or inaction on a policy item.[83]
Policy implementation: Putting an adopted policy into practice.[84]
Policy evaluation: Analyzing implemented policies to determine how well they are meeting objectives. Sometimes followed by changes to improve the policy.[85]

We now describe the way in which the capabilities of AI can affect each of these stages. Of course, it is possible that these stages could ultimately collapse into one simultaneous activity rather than

---

78 Jann, W. and Wegrich, K. (2007)
79 Cairney, P. (2013). While policy does not always follow these stages, this exercise is a useful way of sketching out the various applications of AI to the policy process. These applications can then be applied to other frameworks.
80 Howlett, M. and Ramesh, M. (2003). This model has wide applicability in international practice. It is used in literature about the European Union policy process (Versluis, E. et al. (2011)) and Canadian policy process (Schofield J. & Fershau, J. (2007)), and relates closely to the process used by the U.S. Centers for Disease Control and Prevention ("CDC Policy Process" (2015)). It has been cited as one policy model in a paper about OCED energy policy (Chapman et al., (2016)), and finds grounding in Jann, W. and Wegrich, K.'s (2007) book chapter, which incorporates formulation and adoption into one step. Harold Laswell first proposed the policy cycle in 1951. His model involved more stages, with different names, but the stepwise process has remained similar. Other common frameworks contain modifications, such as collapsing formulation and adoption into one stage, and including a problem identification stage before agenda setting (Jann, W. and Wegrich, K. (2007); Subroto, A. (2011)).
81 Howlett, M.and Ramesh, M. (2003)
82 *Ibid.*
83 *Ibid.*
84 *Ibid.*
85 Howlett, M. and Ramesh, M. (2003); Candler, G. (2014)

discrete steps as described in the above framework. However, for the purposes of analysis, it is useful to look at each of the steps discretely.

A. Agenda setting

Howlett and Ramesh define agenda setting as "the process by which problems come to the attention of governments."[86] Other definitions describe problem identification and agenda setting as the process of recognising a problem and raising its public profile to a point where government officials devote it attention and may consider taking action.[87] Today, these processes occur through various streams of communication between politicians, policymakers, interest groups, and the public.[88] Cobb and Elder divide agenda setting into the governmental (institutional) agenda, containing matters raised by the formal branches of government, and the systemic (public) agenda, comprised of matters raised by the public for action.[89] Groups help elevate or deescalate issues on the political agenda through actions such as providing direct commentary to politicians, protesting, or making political contributions, among other levers. The media may also play a role in framing agenda items and public opinion about them.[90]

AI can affect agenda setting by helping governments aggregate and analyse the interests of the population through various sensors. First, AI will be able to source information from social media platforms to identify problems and gauge public sentiment. Indeed, governments have recently stepped up their efforts to source this type of information. In 2015, the British government announced that it had established contracts with seven companies to monitor information in the public domain on social media platforms.[91] Although the government announcement of the contracts did not outline the specific uses for this information, it would theoretically allow the government to monitor widespread trends in public sentiment, as well as track the public posts of particular individuals. With this type of information, government officials can develop a fuller picture of public opinion as they engage in agenda setting.[92]

AI can contribute to agenda setting by forecasting emerging social and economic conditions,

[86] Howlett, M. and Ramesh, M. (2003)
[87] Jann, W. and Wegrich, K. (2007)
[88] Subroto, A. (2011)
[89] Jann, W. and Wegrich, K. (2007)
[90] Subroto, A. (2011)
[91] Collins, K. (2015); "New Media" (2015)
[92] Collins, K. (2015)

allowing policy solutions to sit one step ahead of problems. Recent academic research has used artificial neural networks to forecast these types of conditions, which are rife with complexity and uncertainty. In one example, neural networks are used to examine public expenses and their correlation with real GDP growth in order to forecast future public expenses by type. This information can help governments react to and prepare for likely expenditures.[93] The greater forecasting abilities of AI make predictive agenda setting more feasible.

B. Policy formulation

The policy formulation stage relies upon policymakers and their staff to help create policy solutions through laws, regulations, and other instruments. It involves the identification and consideration of possible policy approaches and an evaluation of their advantages and disadvantages.[94] In certain contexts, other actors – including experts, non-profit organisations, and international organisations – have increased their roles.[95] With AI, governments can more easily identify individuals, entities, regions, or other factors in the greatest need of assistance or at the highest risk of a particular issue. This information helps AI formulate – and in some cases, deploy – policies specifically designed to address each challenge. Governments may also be able to harness AI to examine and learn from previous policies and their effectiveness.[96]

Consider the example of data-based policy formulation in efforts to combat lead poisoning in Chicago. The typical approach to lead safety involves testing children for lead once exposure is suspected. Using big data, the Chicago Department of Public Health sought to turn this approach on its head by identifying houses likely to contain unsafe lead levels and testing them before families became poisoned. By piloting a model that predicted the likelihood that homes contain harmful lead levels, the department could prioritize houses for lead testing and caution potential homebuyers about high lead areas.[97]

This same type of algorithmic model has functioned in other policy areas, such as identifying students at risk of dropping out of school or street blocks at risk of high criminal activity.[98] Such information lends itself to targeted policy formulation, which helps avoid wasting resources

[93] Magdalena, R. et al. (2015)
[94] Savard, J. F., and Banville, R. (2012)
[95] Jann, W. and Wegrich, K. (2007)
[96] See Russell, S. and Norvig, P. (2014) about AI's ability to learn from prior events.
[97] "Of Prediction and Policy" (2016)
[98] "Data Science for Social Good" (2015); Rudin, C. (2013); Gakrelidz, N. (2017)

unnecessarily.

Finally, in the status quo, policy formulation often occurs in different locations in government depending on the nature of the problem being addressed. Local governments create the majority of policies pertaining specifically to their geographic areas, whereas the national government tends to address issues that cause effects on a national scale.[99] With AI, big data in centralised repositories will help formulate policies to create tailored, localised solutions. In some cases, this may even supplant some local government roles.[100]

C. *Policy decision-making*

Policy decision-making entails the government choosing whether or not to act on a policy matter.[101] It usually follows a formalised procedure, which may involve a period of debate and voting in a chamber of government.[102] During the deliberation period, politicians interact with one another and often strike political bargains. Policy decision-making faces a number of constraints, including many that are resource-based or institutional in nature.[103]

Most likely, this process will change little for now, but will include greater consideration of policy outputs determined by AI. This would look much like the policy adoption process as usual, with the difference being that the proposal policies may be crafted by AI, rather than by human beings. Policymakers could then evaluate these proposals in order to render a decision. Under this model, policymakers might input multiple different parameters for an algorithm to consider, and then compare the best outputs under each set of parameters. This type of human-machine teaming, in which machines and humans collaborate to achieve objectives, has produced fruitful collaborations in fields ranging from gaming to medicine.[104]

D. *Policy implementation*

Policy implementation today involves developing the processes and procedures to translate

[99] "National and Local Government" (n.d.).
[100] Bershidsky, L. (2017); Davies, W. (2016)
[101] Howlett, M. and Ramesh, M. (2003)
[102] Policy Making and Policy Implementation" (n.d.)
[103] Jann, W. and Wegrich, K. (2007)
[104] Holdren, J.P., et al. (2016)

formulated policy into action.[105] It converts policies on paper into practice, designing all steps of who will carry out the process and how.[106] As described by Brewer and DeLeon (1983), implementing policies requires "careful attention," as error can nullify the benefits of the policy and in some cases worsen the condition the policy tries to address.[107] The interests and associations of those tasked with implementing policies can have a strong bearing on outcomes.[108]

Machine learning has the capacity to help determine implementation strategies. For instance, the U.S. Army has used AI for logistics planning for years.[109] However, in the short-term, it is unlikely for machines to take over policy implementation roles (i.e. autonomously determining the best manner for plans to enter into practice), since such tasks tend to involve enormous complexity and would require immense amounts of preliminary data gathering.[110]

One of the first ways that AI will transform policy implementation will entail learning how to target communications to different audiences, rolling out interventions differently across various population subsets. We saw in the "policy formulation" section that AI can help determine which individuals or areas require additional policy focus; with policy implementation, AI can help tailor the scope and style of government interactions with the public to maximise impact.

Nudge theory instructs that the style and manner of government communication with the public can have a large impact on how people react.[111] Observable characteristics such as age, gender, race and ethnicity, and income level all bear on how people respond to government communications and policies. By understanding which types of government communications work most effectively with what audiences, AI can improve the way governments interact and connect with various segments of the population.[112]

For example, after the US healthcare law went into effect in 2010, the U.S. government established a program called Enroll America to identify Americans without health insurance and enrol them in the new healthcare scheme. It used modelling to identify uninsured populations and ran

[105] Savard, J. F. and Banville, R. (2012); "CDC Policy Process" (2015)
[106] "CDC Policy Process" (2015)
[107] Candler, G. (2014)
[108] Savard, J. F., and Banville, R. (2012)
[109] Gleason, D. (1990)
[110] Nowozin, S. (2017)
[111] Howlett, M. and Strassheim, H. (2016)
[112] McCormick, J. (2015)

experiments to test which messages and forms of communication work most effectively for different population subsets.[113] Using demographic information, AI can tailor policy communications to specific audiences. This type of task already occurs in market research, where data collection and analytics drive consumer insights for businesses.[114]

Another way AI will transform policy implementation will entail leveraging social media and other less traditional inputs. In 2015, Las Vegas health department helped implement restaurant sanitation laws and health inspection processes by piloting an app that combs through Twitter posts about food poisoning. The department switched from random restaurant inspections to inspecting restaurants about which people had recently tweeted about food poisoning. AI helped identify the tweets and link them to the restaurants. The tweet-based system increased the rate of health citations by two-thirds, from 9 percent using random searches up to 15 percent of inspections with AI. In total, the system is estimated to have reduced the number of food poisoning incidents by 9,000 during the three-month study period.[115] This type of app demonstrates the ability of AI to serve as a tool for policy implementation.

E. Policy evaluation

Policy evaluation entails assessing whether the objectives set for a policy were achieved[116] and considering whether the policy should be altered or cancelled.[117] While policy evaluation remains a key component of the policy process, it does not always occur as often or as thoroughly as it should. Financial constraints and the availability of skilled evaluators serve as major limiting factors. It can be difficult to assess policies without significant planning and expense.[118] In some policy cycle models, evaluation occurs throughout the entirety of the policy process.[119]

Since policy evaluation is most commonly data-based by its nature, AI offers rich promise for this area. AI will enable more timely policy evaluation than before, without much of the human planning currently required. AI policy assessments in real-time will allow for rapid policy

[113] Stern, S. (2013); Glowacki, M. (2016)
[114] McCormick, J. (2015)
[115] Dubrow, A. and Kautz, H. (2016)
[116] Subroto, A. (2011).
[117] Candler, G. (2014)
[118] Scioli, F. (1979)
[119] "CDC Policy Process" (2015); Jann, W. and Wegrich, K. (2007)

evaluations, as well as policy iterations in response to data-based findings.[120]

Recent research on neural networks has explored AI's implications for policy evaluation. One project examines U.S. government decentralisation from federal to state levels, seeking to determine the impact this had on financial equality among local governments. Using neural networks, the paper found that decentralisation in the late 1980s increased financial inequality among local governments.[121] Other models use neural networks to evaluate the performance of public hospitals and other institutions.[122] This research suggests that AI offers promise for government policy evaluation.

120 Schrater, P. (n.d.); Poole, D.L. and Mackworth, A.K., (2010)
121 Warner, M. E. and Pratt, J. E. (2005)
122 Li, C. and Yu, C. (2013)

V. CHALLENGES

Having identified the opportunities for the use of AI in government, it is important to examine the aspects of AI adoption that policymakers need to get right in order to take full advantage of the technology's potential benefits.

Governments are aware of this technology but are behind the private sector in utilising it. Whilst many governments have acknowledged the importance of the digital age for generating new solutions to their problems, there is no guiding framework for how this tool should be incorporated into current government systems. Central departments everywhere have struggled to articulate a clear roadmap for the future use of MLAs; the White House only appointed its first chief data scientist, D.J. Patil, in February 2015. In an interview following his appointment, Patil discussed the government's commitment to data analysis but voiced concerns over its ability to effectively harness the information available to it due to data silos.

The profusion of data silos is an example of the insular management of data systems within government, and is an important obstacle facing the uptake of AI. UK policymakers have been keen to acknowledge the need for government "to share appropriate information and data across and beyond government boundaries to provide efficient services to citizens, businesses and delivery partners".[123] The foundation for creating such a system is the development of an e-infrastructure,[124] which is defined as "a shared, open (and unbounded), heterogeneous and evolving socio-technical system consisting of a set of IT capabilities and their user, operations and design communities".[125]

In order to build such an environment there are a number of key preconditions that must be met. Firstly, it is required that the law is clear and incentivises data sharing between stakeholders. Secondly, each user (i.e. individual agency) is required to meet a basic technical capacity. Though the design of an e-infrastructure is based on the capability of each agency involved, it is crucial that all parties involved provide strong IT capabilities such that the evolution of that infrastructure can progress efficiently and swiftly. Thirdly, it is necessary that there is sufficient human capacity for adoption to occur, vis-a-vis the talent and human capabilities, as well as the culture, required to embrace the new technology. All of these elements present a challenge to government given the

123 Cabinet Office (2015)
124 Hanseth, O. and Lyytinen, K. (2010)
125 *ibid.*

current status quo, and each will be assessed in turn.

A. Legal capacity

The legal framework is vital for driving open data movements, interoperable data and generally promoting data-driven governance. Privacy concerns have been a key aspect of this debate. The International Telecommunications Union has defined privacy as the "right of individuals to control or influence what information related to them may be disclosed".[126] AI's reliance on data means that any government wishing to use these tools must assess privacy legislation to ensure it is compatible with their work.

Understanding who owns the data and what, legally speaking, can be done with it is key to unlocking its potential. Data ownership is a complex and politically fraught field. It is imperative to tackle public concerns about data use and provide a clear framework for data privacy laws that benefit both the individual and society more broadly. Understanding the nature of the threat is crucial to our attempts at tackling it. A recent White House report asserted that "[t]hreats to privacy stem from the deliberate or inadvertent disclosure of collected or derived individual data, the misuse of the data, and the fact that derived data may be inaccurate or false".[127]

The threat to privacy is not a new one but AI brings its own set of issues. In the modern era an individual does not necessarily own their own personal data; it could have been collected using public sensors such as cameras or on public domains such as Twitter. Data analytics complicates this issue as it can use data sources without authorisation and produce an unknown output. It has been noted that "[t]hose analyses sometimes yield valid conclusions that the individual would not want disclosed. Worse yet, the analyses can produce false positives or false negatives".[128]

Unlike the private sector, government does not have to compete with other institutions in setting a precedent for its data practices. It is also distinct in the impetus derived from its responsibility to ensure national security. Both factors recognise that the public sector faces specific challenges concerning privacy rights and as such these warrant specific solutions and legislation. This issue is lent further subtlety when one considers the increasingly blurred lines between public and private data. Both sectors can potentially utilise the same data sources and algorithms, making the

126 Akerkar, R. (2014)
127 White House Report (2014)
128 Executive Office of the President (2014)

negotiation of privacy laws difficult to manage. Data is produced from a variety of sources, some of which are state controlled, and others which are not. Governments might, for example, wish to integrate data from private companies in order to access information not currently available to them through publically available sources.

An example of this would be the US National Security Letter (NSL) which grants the FBI the right to subpoena customer records held by banks, telephone companies, internet service providers, etc.[129] The exchange of data is not unidirectional, however – for example, companies such as BUPA can purchase medical records without personal consent.[130] Usually provided as part of large datasets, this data is meant to be anonymised but it raises important questions about the ethical and legal nature of such practices. The law will need to take these concerns into account and efforts must be made to make sure governments do not over-assert their rights. Whilst ideas about privacy change generationally, it is crucial that policy both serves the needs of government *and* citizens.

B. *Technical capacity*

Government silos are exacerbated by a lack of processing power and unequal data quality across agencies. Big data is the raw material of machine learning and large scale processing power is needed for its use. According to Intel, "[h]aving the right IT infrastructure, including the server hardware, software, networking, storage, and data management technologies, are all essential parts of the puzzle when it comes to Big Data". Whilst central governments have begun to build the infrastructural requirements necessary, data storage needs to evolve from its current SQL format to diverse formats such as MapReduce or Bulk Synchronous Parallel (BSP), which allow structured and unstructured data to run on parallel processing networks.

Not all data is created equal and governments do not always have the best quality of data at their disposal. Good quality data is accurate, complete and timely. With a growing reliance on data to support government decisions, the data public institutions accrue must be standardised. The public sector has as yet been unable to consistently achieve this, thus limiting the production of quality information. Research suggests that analysts spend around 90 percent of their time cleaning data and "data, especially government data, is often provided in non-machine readable or non-standardised formats requiring manual re-entry".[131] Another limitation is the burden of maintaining

129 Ackerman, S. (2016)
130 Sparrow, A. and Mulholland, H. (2011)
131 Bulger et al (2014)

current hardware and software capabilities (e.g. sensors such as CCTV must be maintained and improved), so as to provide governments with the high quality data necessary to realise the full value of these algorithmic techniques.

C. Human capacity

The technical and legal obstacles faced by government only tell half of the story. Human capacity is also an important factor when considering how to generate an integrated e-infrastructure. Capturing the value of AI will involve more than just recognition and aspiration; past the vision required by public sector leaders to seek out opportunities for MLAs, there is a real and significant challenge to meet with regards to talent and capability.[132]

Current labour trends show that governments have placed building human capacity, specifically in data science, high on their agenda. Whilst the public sector attempts to build capacity in this area it faces a lag time in terms of in-house data specialists. "Numerous studies have estimated the potential jobs gap for data scientists and related disciplines – and put the figure for the shortage of staff with deep analytical talent in the United States in the low hundreds of thousands by the early 2020s."[133] Without the government workers who really understand this technology and the benefits it can bring, development will remain slow and costs high as governments are forced to rely on outsourcing to technical specialists/consultants.

The culture within government does not always enable departments to fully exploit the potential of AI. There are time lags and strongly enforced hierarchies which can mean innovation is slow to come. This does not directly constrain the uptake of AI but can shape the way opportunities are assessed and developed. Employee buy-in is crucial to securing support for the necessary structural and procedural changes departments might have to implement in order to facilitate the adoption of this tool.

132 Yiu, C. (2012)
133 Policy Exchange (2012)

VI. RISKS

A. What if we do nothing?

What the proper role of government should be is contingent on what people believe constitutes the 'good society', as well as their general confidence in government as a means of improving society.[134] We argue that if government does not keep up with technology it will provide lower quality outcomes relative to the private sector, and this will undermine the legitimacy of government as the central provider of solutions to societal problems. By falling behind on the AI curve, governments will have serious shortcomings when it comes to solving the large-scale problems they oversee. If corporations are able to employ AI to make better decisions than governments, the private sector will be legitimated to encroach upon government functions.

i. The output v input model of legitimacy

Legitimacy is at the core of all citizen-state relations. It determines how power is used by a state and can be understood as "an acceptance of authority by both elite and non-elite groups,, although not all citizens are equally able to confer legitimacy".[135] Whilst many models have been proposed to describe state legitimacy, in this report we draw upon the input-output model. This has gained traction in recent years amongst political scientists due to what some governance theorists have described as a shift from input-oriented forms of democracy to output-oriented ones. This shift has been ascribed to changing notions of accountability stemming from the New Public Management reforms, which sees citizens play a greater role in ensuring accountability and in monitoring policy outcomes.[136] On the one hand, as service delivery mechanisms become more complex, democratic participation is occurring within the institutional arrangement of bureaucracies. On the other, the state-citizen interface is gradually emphasising delivery and output rather than representation.

As such, political legitimacy depends at least as much on the quality of government than on the capacity of electoral systems to create effective representation."[137] A ruling regime will therefore need to strengthen belief in its ability to deliver effective solutions to problems. Given that, as Levi et al. have demonstrated, compliance rates decline if a government is perceived as not being

[134] Chalmers, J.(1959)
[135] McCullough, A. (2015)
[136] Pierre, J., Røiseland, A. and Gustavesen, A. (2011)
[137] Rothstein, B (2009)

competent enough to deliver on its promises and solve problems,[138] we argue that government will lose legitimacy if it is perceived to deliver outcomes inferior to some other entity.[139]

ii. Declining outcomes

The scope of government today is broad, but includes operations, service delivery and policymaking. In the process of performing these functions, governments collect vast amounts of data on their citizens. These large and diverse datasets are difficult to organise, manage and extract insights from, yet governments need to do so in ways that will benefit the general public and facilitate government transparency. Some analysts estimate that governments could act as catalysts for over $3 trillion in economic value if they were to digitise information, disseminate public datasets and apply analytics to improve decision-making.[140]

But governments don't have a monopoly on information. Companies like Google and Facebook amass vast amounts of consumer data, which they utilise to improve their service delivery, product quality, and marketing.[141] Governments are lagging behind on this front. In order to fulfil their core functions, governments need to get better at collecting and consolidating data, and they need the in-house expertise to analyse and utilise the data they collect.

Access to knowledge and data

Citizens trust their government to look at information at an aggregate level. Governments make decisions about where to spend public money and how to structure policies because they have access to more information than individual citizens. AI matters to governments because AI technology has the power to democratise information, thereby dispersing decision-making capabilities. Given that access to information leads to better decision-making, what will happen to governments if they no longer possess more information about their citizens than other entities?

Whilst some might be sceptical about whether a single technology has the power to change deep-seated societal structures, we would argue that history proves otherwise. Take the example of the printing press, which scholars have shown led to the decline of the role of the Catholic Church in

[138] Levi, M, Sacks, A, and Tyler, T (2009)
[139] There are many ways of understanding legitimacy, and we are using one of many models. For a longer discussion on what legitimacy means and where state legitimacy is at today, please see CPI's global discussion paper What drives legitimacy?
[140] Moreno, H. (2014)
[141] Waddell, K. (2015)

Europe and paved the way for the emergence of modern states. According to the American historian Elizabeth Eisenstein, the Renaissance and Reformation could not have occurred without the printing press, as it was instrumental in producing the volume and diffusion of text necessary to allow diversity of thought.[142] The printing press allowed for identical words and images to be reproduced and then viewed simultaneously by scattered readers. The ubiquity of print loosened the Church's authority by undermining the faith's unifying myths. Notably, Martin Luther's message was not all that different from that of the Czech priest Jan Hus, who had demanded to change some of the church's core principles, but was condemned to death and burned at the stake in 1416.[143] But Luther's writings, produced after the advent of printing, proliferated and reached a large number of readers, representing more than a third of the total number of books sold in Germany between 1518 and 1525.[144] They produced a snowball effect: people felt encouraged to speak out publicly against the Church, and they were followed by others, and then yet others.[145]

There is little doubt that this new 'information technology' was a catalyst for the Reformation, and within a relatively short period, the Catholic Church lost its hegemony over public opinion. It is not inconceivable that the rise of AI could have similar consequences for governments today, should they fail to capitalise on its opportunities and guard themselves against its risks.

For example, following Richard and Daniel Susskind's argument, it is not inconceivable that in a knowledge-based society where AI is ubiquitous, white-collar professions – including government and the public sector – will undergo radical transformation.[146] In the Susskinds' argument, the move from a print-based industrial society to a technology-based internet one has major consequences for professional 'gatekeepers', because specialist knowledge previously confined to bookshelves or professional standards bodies, and kept opaque, becomes directly available to non-specialists online. As the internet changes the way in which expertise is produced and distributed, and new, capable machines emerge that do not require users to have formal knowledge or skills to access this expertise, the role played by professionals becomes less clear. In the process, the 'grand bargain' - a social contract between society and the professions which elevates and shields professional groups through legislative and regulatory procedures – breaks down. Whereas previously, professionals were seen as the rightful gatekeepers of knowledge, in the new society,

142 Guthrie, J. (1982)
143 Kertcher, Z. and Margalit, A. (2006)
144 *ibid.*
145 Rubin, J., (2012)
146 Susskind, R., and Susskind, D. (2015)

their trustworthiness is no longer taken for granted and their very utility is questioned.

It is possible to see the nascent beginnings of this phenomenon today, in certain fields of knowledge previously presided over by government departments. Take for example the health and education sectors. Websites like the WebMD network provide an abundance of guidance on symptoms and treatment, and there are more unique visits to the site each month (190 million) than to all the doctors working in the US. Online communities like PatientsLikeMe have around 300,000 people who share their conditions with other people (at the moment, around 2300 conditions), and swap experiences and treatments.[147]

In education, schools where technology is central are experimenting with completely different teaching systems. At Rocketship Education, a network of nine charter schools in California, students spend a quarter of their day using an online platform called the "Learning Lab" without the need for an actual teacher. There are massive online platforms that provide free educational content. Khan Academy has a free collection of 5,500 instructional videos and an effective attendance higher than the total primary and secondary schools of England (10 million unique visitors each month since 2014). Along with other online platforms like TED-Ed, and YouTube EDU, they have contributed to flipping the classroom experience – educators are assigning lectures to be watched at home and using classroom time for 'homework'. Parents are also using online content for homeschooling, a practice that has doubled between 1999 and 2012.[148]

For governments, such potentially profound changes in access to data and knowledge by the public represent both opportunity and risk. On the one hand, governments have already been able to make use of similar platforms to those described above, to solve societal problems more efficiently and effectively – take, for example, NHS Choices which, like WebMD, has been hugely successful. On the other hand, as the Susskinds' argument shows, there is a significant threat involved in the democratisation of knowledge, whereby increased transparency of the work of professionals results in public disillusionment with government performance. Governments need to be alert to such prospects, and to the changing ways in which publics will hold governments to account in a digital, AI-facilitated age. Whereas previously, professionals engendered trust from lay people by cultivating a perception of ethical and moral superiority, open access websites like Khan Academy are trusted by publics because they are seen to operate reliably and receive excellent reviews from

[147] Susskind, R., and Susskind, D. (2015), p. 52
[148] Susskind, R., and Susskind, D. (2015), p. 57

members of *the public themselves*. As these fundamental shifts in public accountability and trust continue, governments will need to be both reactive and anticipatory in their outlooks, if they are to maintain their legitimacy going forward.

Capacity to use knowledge and data

Governments have been collecting vast amounts of data on their citizens long before the private sector made the concept of big data popular. But most governments aren't actually making use of these vast troves of data. Not only is public sector data collection haphazard (for example, data is split between specific departmental systems and often cannot be compared across the entire organisation), but many agencies lack the necessary in-house talent and infrastructure to maintain and analyse all of the information they possess. According to one report, public-sector data analysts say that they spend almost half of their time collecting and organising data, but less than a third of their time gleaning actionable insights from it.[149] A lack of technical expertise in organising and analysing their data has also meant that governments have had to pay for assistance from the private sector. We argue that not only are these missed opportunities for governments, but that there are potentially serious risks involved should government lethargy in the uptake of AI continue.

We are already seeing government pulling out of offering services in certain areas and/or outsourcing service provision to private companies. For example, in 2016, a suburb of Tampa in Florida experimented with replacing two bus lines and subsidising Uber rides for its locals instead.[150] At $40,000 a year, the programme will be about a quarter of the cost of the two bus lines it is replacing. Lyft (Uber's main competitor) has also been helping a dozen transit agencies apply for federal grants to cover the app's ride fares in a move that will effectively make Lyft drivers part of the public transportation system.

Although these projects are at the moment fairly small-scale (the county in Florida will only be spending 16 percent of its $3.5 million federal grant to improve public transportation on subsidising Uber and Lyft rides to two train stations), if ride-sharing apps can provide better outcomes for the population and do so at an equal or lower cost than the public transit system, we can expect this trend to continue. In fact, by some estimates, there are twice as many people doing

149 Latham, E. (2016)
150 Brustein, J. (2016)

government work under contract than there are government workers."[151] According to the Washington Post columnist Steven Pearlstein, "Federal contracting grew from about $200 billion in 2000 to about $550 billion in 2011 before falling back to $450 billion last year". Sixty percent of that was for services.

Yet there are, and should be, serious concerns about what this kind of private sector service provision might mean. Three concerns that we have identified are the regulatory and labour issues surrounding private sector companies' service provision; the ethical concerns surrounding private sector use of government data about its citizens; and the technological deskilling of government staff, as data management is increasingly outsourced. Finally, of larger concern is the way in which transferring capacity to the private sector will eventually undermine government legitimacy in being able to provide for its citizens.

Concerns about unfair labour practices at Uber have been well documented. Whilst the taxi business used to be highly regulated, Uber and other e-hailing taxi apps have leveraged private-transport matching to provide cheaper, more effective cab options. We suggest that as technology revolutionises traditional industries, governments will need to develop faster ways of responding to such changes, in order to keep protecting their citizens. This will be even more imperative in industries like the taxi business, where governments are themselves relying on private sector service provision, and will be judged even more strongly on whether they are able to ensure a robust regulatory environment.

Of even greater concern with regard to private sector service provision is the use (and potential misuse) of citizens' data by private companies. For example, the recent partnership between the Royal Free London NHS Foundation Trust and the AI company Google DeepMind has resulted in some controversy. The deal gave DeepMind permission to process 1.6 million Royal Free Hospital patients' records from November 2015 to November 2016.[152] Under this deal, DeepMind uses patient data to deliver early warning signals as part of a five-year contract using their acute kidney injury (AKI) alerting app, Streams. Earlier this year, DeepMind signed up Imperial College Healthcare NHS Trust for a similar deal, and the company also has artificial research partnerships with two other London NHS trusts.[153] Mustafa Suleyman, the co-founder and head of applied AI at DeepMind, says the company has a big future in the NHS, one that goes beyond apps and a

[151] Pearlstein, S. (2014)
[152] Shead, S. (2017)
[153] Digital Health (2017)

handful of research partnerships.

Whilst the app has proved successful (apparently saving nurses up to two hours every day, and giving them more face-to-face time with patients),[154] the deal has been marred by privacy concerns, data access issues and an ongoing Information Commissioner's Office investigation over whether it was legal under the Data Protection Act. Following the publication of an article entitled "Google Deepmind and healthcare in an age of algorithms" in *Health and Technology* in March 2017,[155] many ethics experts believe that the deal was hastily concluded with little attention paid to how data was being shared. Since access to data is so crucial to AI firms, a deal this favourable for DeepMind at the dawn of the health analytics industry runs the risk of giving it a monopolistic advantage in the future. Further, it highlights an asymmetry in government and private sector approaches to data, where for the former, data management is viewed as onerous, whilst for the latter, it is seen as trade secrets and key to a firm's competitive advantage.[156] With government-acquired health (and other) data, "we need to be sure we aren't, in effect, giving oxygen away for free to a private company that will start to sell it back to us."[157]

Thirdly, outsourcing to the private sector is likely to be a continuing trend in the wake of a limited talent pool and budgetary constraints. Unless action is taken, it will be difficult for governments to recruit and keep the talent required to fully exploit AI's capabilities, given caps on government employee numbers and salaries that are two to three times lower than that achievable in the private sector. More to the point, a combination of increased outsourcing and poor internal capabilities can lead to a number of problems. The Edward Snowden leaks about extensive spying by the NSA and the botched Obamacare rollout are both manifestations of this. The Obamacare website is a particularly striking example: when the private contractors (the tech conglomerate CGI Federal) failed to perform as promised, the IT staff at the Centers for Medicare and Medicaid Services were unable to step in. They did not have the necessary training to either manage the contractors or oversee the final integration of the new system.[158]

The concerns highlighted above demonstrate the need for governments to address the risks

154 Burgess, M. (2017)
155 Powles, J. and Hodson, H. (2017)
156 Brustein, J. (2016)

157 Wong, J. (2017)
158 Lipton, E. et al (2013)

emerging from a failure to develop the internal capabilities required to take advantage of AI opportunities. Yet, even if governments manage to successfully react to the regulatory, ethical and human capacity demands discussed above, the even bigger risk – outlined at the start of this section – remains: in a world where the private sector is better able to serve the needs of the public, what role remains for governments, and where will they find their legitimacy? The balance in society between state and corporate power has long been complex, but a scenario in which public service provision is increasingly taken over by a more capable private sector will mark a substantial shift. The question then becomes: will these private corporations serve the *public*,[159] and what will governments be able to do about it?

B. *What if we get this wrong?*

In the previous section, we discussed the potential consequences of a scenario in which government falls behind the private sector and parts of the general public in leveraging AI capabilities, eventually resulting in a loss of government legitimacy and a fracturing of the social contract between government and society.

In this section, we propose an alternative risk scenario: one in which governments *do* leverage AI, but do so in ways that could be perceived as detrimental to the social good. Within this scenario, we explore the potential for AI uptake to produce systematised inequity, spiralling costs, and abuse of government power. Whilst these kinds of ethical concerns permeate the functioning of governments irrespective of technology use, the introduction of AI is certain to exacerbate particular issues and bring forth its own, new challenges. The consequences of failing to manage these ethical risks are high.

i. *Systematised inequity*

Algorithms are beginning to be used by governments as arbiters, to make decisions like who receives welfare benefits or which individuals can be granted parole. Earlier in this report, we described the hidden biases that inhere in human decision-making, and how AI can help to make decision-making fairer and more consistent. However, that AI use will result in better decision-making is not a given. There are a number of problems associated with the use of AI in decision-making, including that bias is reintroduced through human interaction with algorithms and/or biased input data; and that the values and weightings required by MLAs ultimately rely on

[159] Among the many issues that would need to be resolved are access to rides for people without smartphones, services for disabled riders, and the conversion of public sector employees into contractors.

subjective human value judgments, which become hard-coded into systems. Arguably, these pose a bigger threat than poor human decision-making in the absence of AI, because the results are not just poor outcomes but systematised inequity.

Whilst in theory AI is noise-free and neutral, in practice it is often neither, due to the data that is input. Data can be noisy due to spurious readings, measurement error, or background data, resulting in randomly inconsistent and inaccurate outputs. Even unsupervised MLAs are dependent on the subjective decisions made by the humans who select and tag the data used to train them. An example of this is the filtering of objectionable content in social media. To train an algorithm to identify such content, a human has to flag a series of cases that they find sensitive for reasons that may be based on personal or cultural norms. As such, AI decision-making becomes no better than human decision-making, but may mask biases behind a veil of assumed objectivity. This reinforces the need to audit the data used to train an algorithm with the same rigour as the algorithm itself.

Even more worrying is the fact that all input data is derived from an inherently biased status quo. Since AI 'learns' from its input data, there is the risk that it systematises this biased reality, producing *consistently biased* outputs. For example, many hiring decisions in the private sector are being made algorithmically.[160] Humans have been proven to show bias towards people from their socioeconomic background and sex,[161] as has been evidenced by numerous experiments using 'blind' employment processes.[162] Yet, it has been shown that algorithmic hiring processes also discriminate against minority groups because long commuting times are correlated with high staff turnover and this, in turn, is correlated with minority group status. Thus, even when an algorithm has been deliberately designed to omit contentious variables (e.g. race), it can be unconsciously biased. Despite deliberately choosing the data used to train algorithms, the correlations that algorithms work out to fulfil their functions can come from sources of which their programmers are unaware. Identifying these problems will often only be possible after they have informed potentially damaging decision-making processes. There are still no standardised practices for ensuring that algorithms do not introduce new biases, and in judging algorithms, governments tend to have limited information, relying on the subjective interpretations of data scientists.

160 Kneese, T., Rosenblat, A. and Boyd, D. (2014)
161 Altonji, J. and Blank, R. (1999)
162 Goldin, C. and Rouse, C. (2000)

Even before the data input stage, there may be significant ethical issues involved in the use of AI. Since MLAs rely on concrete definitions of values, it is necessary to pre-specify societal values into algorithms' codes. Yet, in any given society there are often irreconcilable conflicts about values, particularly surrounding fairness and equity. One example is the distinction between equality of treatment and equality of outcome.[163] If an AI is programmed to be fair according to the definition of equality of treatment, we will get different outputs than if we program according to equality of outcome. For example, in the case of college acceptance, it has been shown that SAT verbal questions disadvantage African-American students.[164] According to these findings, to achieve fairness of outcome the AI should account for different scores across subgroups and treat students differently. Whilst this is not a new problem for policymakers, with the use of AI they will be forced to agree on a definition of fairness on which to base the model and benchmark performance. In strategic decisions, definitions of fairness may have particularly long-term policy consequences.

Further, even where policymakers are able to agree on value definitions, there remains the issue of weighting different desirable values, such as fairness and utility, or weighting the various characteristics that alter and define an outcome. For example, an MLA that identifies cancerous cells in an image must be programmed with confidence intervals. For a programmer without medical knowledge, a confidence interval of 5 or 10 percent might be acceptable (in practice this is tested against human accuracy rates); for a doctor, the only acceptable diagnostics error term is zero, as the professional consequences of making a mistake of this kind could be devastating. If algorithms are statistically more accurate than humans at making such diagnoses, they should also be made accountable when they make mistakes, as doctors are, with this accountability being extend to the humans responsible for the algorithms (e.g. the programmer or commissioner).

Arguably, then, policymakers and/or AI programmers are tasked with even greater responsibilities than they currently face. Whereas presently, overall policy decisions are set centrally but decision-making about individual cases is dispersed across a multitude of human decision-makers, in a system where AI undertakes all decision-making, the central function assumes much greater powers. AI can potentially systematise a biased reality. Further, there is a danger that because it offers the potential for cost savings, there will be a temptation to use AI to scale decision-making

163 Friedler, S. et al. (2016)
164 Santelices, M. and Wilson, M. (2010)

processes.[165] Such scaling up will result in the systematic classification of certain types of individuals within the system. For operational-level decisions, this may be less of a concern, as we may value consistency more than diversity. At the strategic level, however, it can result in consistently excluding or marginalising entire social groups.

ii. Spiralling costs

Whilst cost reduction is one of the central reasons cited in framing the decision to introduce AI, a major risk to government is the hidden costs that AI adoption brings. The assumption that AI delivers cost-savings is not unambiguously true; the costs of implementation are often indirect, difficult to calculate, and therefore underestimated. For example, take the US job-matching search engine for veterans. The search engine costs $5 million per year, but is only used by a couple of hundred veterans. Across the board, Tim O'Reilly found that there was no routine process for assessing the cost-effectiveness of government websites.[166] Government approaches to AI appear to be pervaded by an assumption of cost-savings, leading them to discount the need to measure and plan for costs. Development and implementation expenses are rarely taken into account, particularly lifecycle costs like data and model management, oversight mechanisms and auditing expenses. Further, the risk is that once AI systems are implemented, their removal will be difficult due to their embedding in complex processes and a concomitant reduction in the human capacity required to fulfil the role they once performed. If this is the case, governments will be locked into particular AI systems, with the potential for spiralling costs going forwards.

iii. Abuse of government power

A third, and more ominous, ethical concern is the potential for government misuse of citizens' data. For governments to take full advantage of the benefits of AI, it will be necessary to share and centralise data across all of government. Yet, the risks associated with this are huge. On the one hand, the requirement for systems to be robust (for example, to hackers) will be even more acute than it is at present. On the other, there is also the need to be wary that government itself does not overextend its powers, turning into an all-seeing "Big Brother" state, monitoring and manipulating the minutiae of its citizens' lives.

Even where governments do not move into new functional areas of control, the increased automation of decision-making through AI use may be perceived by the public as dangerous. Fears

165 Boyd, D., Levy, K., and Marwick, A. (2014)
166 Adler, T. (2014)

of faceless decision-making may be well-founded should algorithmic complexity mean that decision-making processes become inscrutable. If the processes by which decisions are made are embedded in mathematical formulae that are impenetrable to the average citizen, transparency and accountability will be diminished. Further, these changes will most likely affect the least advantaged groups in society, who are both more likely to be subject to the outcomes of these decisions and less likely to understand the processes involved in producing them. This reduced level of understanding will result in a diminished possibility for raising objections, shifting the balance of power away from publics and towards governments. In contrast to the scenario in the previous section, then, government may become stronger, but it will be less liable to being held to account, arguably to the detriment of the social good.

How scrutable an MLA is depends on its complexity. As a broad rule of thumb, the greater the sophistication of an algorithm, the lower the level of human intervention and understanding. Algorithms such as Bayesian Networks (used to detect probabilistic relationships) and decision trees (used for classification) are relatively straightforward and can be scrutinised and evaluated for identification of structural problems.[167] However, more sophisticated algorithms such as Deep Neural Networks are not equally transparent to inspection; their learning process is so independent of human input that it is almost impossible to determine what identification parameters they utilise to make decisions.[168] In such cases, if even the programmers who create these machines do not know how they work, who should be held accountable for their results and mistakes?

The view from some quarters is that the solution is not to try to understand the MLA, whose function is merely to process data. Instead, the answer is to scrutinise the data used to train the algorithm. Alongside another school of thought, we disagree with this view, arguing that tackling the inability to generate effective models for inspection of these sophisticated algorithms is the responsibility of data scientists, whose contribution should not be limited to technical work. Their role could be extended to include, for example, the development of reporting formats and auditing techniques such as algorithmic audits, which aim to discover what sets of features result in algorithms making particular decisions.[169]

Ultimately, we argue that mitigating the ethical risks involved in AI use requires that people – and

---

[167] Hastie, T., Tibshirani, R. and Friedman, J. (2017)
[168] Bostrom, M. (2015)
[169] Rieder, B. (2015)

not technology – be in control: if AI is to be successfully implemented by governments, expertise and transparency is required at all levels. Governments need to ensure that they have the personnel to interact meaningfully with AI; and the outcomes of AI decision-makers need to be scrutable by the public. Finally, when holding AI to account, it will be necessary to have its human creators visible and held ultimately responsible for decisions, to avoid the deferral of agency to machines if and when things go wrong.

C. Will AI fundamentally change the way government looks, thinks and acts?

If the short-term consequences of AI are already challenging to predict the longer-term effects are even more unclear. The use of AI in government might well lead us to a future where governments simply operate somewhat more efficiently and somewhat more effectively. Or it might have more transformational effects and fundamentally change the way governments look, think and act. Which way the journey goes is, of course, not only determined by the inherent potential of the technology but as much, if not more, by the social and political choices we make. Exploring such longer-term futures isn't useful because it yields precise predictions (which it does not) but because it illuminates which decisions we will need to take and what their consequences might be.

What might the consequences of such a transformative future be? Given that the purpose of this particular report is not to explore such longer-term consequences exhaustively we will take a cursory look at three areas that might change if AI was to transform government: the balance between evidence and values, the organisational structure of public administrations and the speed of decision-making. Today, the best policymakers spend much of their time debating questions of fact to ensure policy is based on the best possible evidence of what works. The problem is that human comprehension and ability to absorb information is limited, so at best we only get a partial picture.

In the future AI might be able to absorb far more information far more quickly than any human about the most effective approaches in education, health, justice or any other policy area. If that was the case policymakers could focus their attention on defining the normative goals to be achieved (and the boundary conditions) rather than spend time debating question of fact. It may turn out that this is what democratic institutions like parliaments are fundamentally good at. At the same time this process would force us to spell out all value judgements and trade-offs in painfully transparent ways – something which might well lead to paralysis and gridlock.

The organisational structure of most public administrations is nearly identical. Almost all governments organise themselves in departments formed around supposedly distinct policy areas – energy, transportation, trade, work and pensions, etc. This is, of course, for the very sensible reason that organisations, and the humans who work within them, need a degree of focus. Humans are also limited in the amount and diversity of information they can usefully process. Already today, however, most difficult problems governments face don't align with neat organisational responsibilities but cut across many areas. If governments relied on AI more heavily they might find themselves compelled to realign their structures such that they align around the problem rather than around arbitrary "policy areas".

Finally, the interface between citizen and state remains crude to this day. In most democracies, citizens are invited to express their preferences once every few years through the ballot box. Otherwise, their ability to influence the nature public services is limited. A government that fully leverages AI would know – and understand – much more about its citizens than in any previous period in history. Citizens would be extraordinarily "legible" to the state, to use James C. Scott's terminology. Government could build a fine-grained representation of each citizens and their preferences and adapt public policy as a result, in real-time if necessary. The implications for traditional democratic models could be profound.

The reflections above are speculative, cursory and incomplete, yet likely worth exploring further since we will only be able to shape the outcomes and create a desirable future if we understanding what's at stake.

VII. RECOMMENDATIONS

Government interest in AI is gaining momentum and this technology has huge potential to transform public services if implemented carefully.

The following recommendations are directed at officials in all government departments, and represent our views on what is required in the short term to facilitate the responsible uptake of AI for the improvement of existing government functionality. No doubt, as AI's use becomes both deeper and more widespread, such recommendations will require refinement in order to address new opportunities and new challenges, as well as the changing role of government itself. In writing to government officials in the broadest of senses, we consider the capacity heterogeneity between different countries defined in terms of people, tools and systems, and also the capacity heterogeneity between government agencies. Consequently, we divide our recommendations into three elements:

- ✓ **Define needs:** Best practices for identifying departmental need
- ✓ **Build capacity:** Human and technical building blocks required for the uptake of AI
- ✓ **Adapt structures**: Adaptations required to existing cultural, regulatory and legislative environments

A. Define needs

Whilst AI is a transformative technology that will drive change in the way governments operate for years to come, its implementation should be iterative. No government department will be starting from scratch: all government departments have data capabilities and technologies that they can build on.

Identify departmental needs not met by traditional approaches: In our analysis, we found that successful applications started with a policy problem that could be solved more efficiently and effectively with AI. Instead of starting with the technology, these applications first identified important departmental needs that were not being met by traditional approaches and then asked, "How can AI help?".

Use inaugural AI applications to identify new ones: We also found that once AI-driven solutions were developed for these inaugural applications, departments discovered other policy problems that could be solved with similar solutions that built on prior investments. For example, the City of Chicago scaled up its CrimeScan software to the city-level, introducing it as part of its SmartData platform that will be used for real-time predictive analytics and decision-making in areas such as rodent control, preventing STIs and emergency response.[170]

B. Build capacity

i. Human talent

A few years ago, Harvard Business Review placed data analyst as "the sexiest job in the 21st century".[171] The private and public sector are competing for an extremely scarce talent pool – the data talent gap was one of the most pressing issues discussed at the 2015 World Economic Forum – and the private sector is currently winning this competition.[172]

Governments can do three things to address the issue:

1. Increase the size of the talent pool: Being at the forefront of training data professionals will derive long term economic benefits for countries as demand for these services increases worldwide. The US has recognised this pressing need and is working to maintain its status as the world's top data scientist talent source. The US Department of Homeland Security set up a programme with universities including Carnegie Mellon University and the University of Maryland to train a pipeline of approximately 30,000 data experts. Adopting similar partnership initiatives between the State and universities is a good first step, however, it is likely to be insufficient in the short term as it will take several years for the benefits to be fully realised. To plug immediate needs, governments must be open to recruit globally without discriminating based on citizenship status.

2. Formalise a data scientist career path within the public sector, and provide additional data analytics training for current public service staff: New recruits are assigned into ordinary public service positions that do not fully focus on data analytics such as Government Operational Research Service (GORS), Government Statistical Service (GSS), or Government Social Research Service.[173] Therefore, the attractiveness of a data scientist job within the public sector is diminished compared to the private sector where specialisation is often a necessity. Hence, it is crucial for

[170] Neill, D. (2013)
[171] Davenport, T. and Patil, D. (2012)
[172] Steele, P. (2015)
[173] Yiu, C. (2012)

governments to formalise a data scientist career path, not only for top-level posts (e.g. chief technology officer) but also at analyst level.

Talent can also be developed internally through additional training schemes. There are significant advantages that come with building talent from the inside due to an awareness of the organisational setting and a deep knowledge of public sector operations.[174] In 2014, the UK Government spent £150,000 on 150 public servants for an open data training scheme. Such schemes have been proven to increase employee satisfaction and go some way towards harnessing the talent already in government. Facilitating the development of specialised roles and training public sector workers should go some way to creating the necessary capacity in government to harness this tool. However, the downside is that it opens up opportunities for employees to move onto new jobs outside of the public sector.

3. Assert the benefits of public sector work and target initiatives to interested parties: Government is unlikely to be able to match the likes of Google and Facebook in terms of wage rates but it should focus on advertising the benefits of working in the public sector. Indicators such as hours worked, gender equity in the workplace and worker satisfaction are areas where government is best positioned to compete with the private sector,[175] and it should focus on broadcasting these to a broader audience. Another potential technique for retaining high-value staff is to seek out university partnerships that emphasise the acquisition of data skills for the public sector.

For example, the University of Chicago has set up a new programme, *Data for Social Good*, which promotes itself as a degree specifically tailored to public sector needs. “Program fellows work in teams with front-line decision-makers on high-impact data analysis problems and are paired with a full-time academic or industry mentor who serves as a technical advisor and project lead”.[176] Rayid Ghani, Obama´s former chief data scientist for the 2012 campaign, heads the fellowship. Such programmes help attract talented graduates who want to work in the public sector, and governments should invest time and money in developing these types of initiatives.

[174] Steele, P. (2015)
[175] Rogers, S. (2012)
[176] Sommerfeld, C. (2015)

ii. Technical capacity

Building a technology ecosystem for utilising this tool is key. We can consider technical capacity as internally and externally powered. Both elements will be addressed in turn.

Governments must invest in their processing power: In order to build the infrastructural requirements for the use of AI, it will be necessary to improve current processing capacity. Given current trends which suggest data processing capabilities will shift to the cloud, governments are set to benefit from the evolution in technology which will make computing services cheaper and better.

To generate immediate value, governments should start by using the data collection devices that they already own: The public sector has proprietorship of sensors such as satellites, video cameras and telecommunication antennas. Making use of these existing devices, governments could easily adopt opportunities such as face recognition, adding the requisite technological upgrades alongside scheduled updates to their IT infrastructure.

C. Adapt structures

i. Culture

Bureaucratic obstacles are a major challenge for the integration of new technologies in the public sector. Further, interdepartmental differences in terms of perceived need and capacity mean that not everyone will show interest in AI.

Find a champion: Government needs a leader who has the vision for the advancement of data-driven governance and technological innovation. Many cities in developing countries have created the role of the chief data officer (CDO), who is in charge of infusing technological innovation into various departments within government and improving IT capability in order to meet strategic goals and needs.[177] However, creating a new role is not sufficient: the CDO must find agencies interested in this technology and work with them to capitalise on early gains.

Secure early wins: A progressive and open mind-set will aid the adoption of data-driven governance, but in the event of stakeholder resistance, the best way to change minds is to show

[177] Stephens et al (2015). Footnote to be updated.

results. Governments should start by identifying key aspects to target within a broader goal, setting baselines and targets in both standard performance metrics (e.g. cost reduction) and under-recognised metrics like reduced human error and data standardisation. This will demonstrate progress and allow for the monitoring of continuous improvement.

ii. Legal framework

Regulation is required to set bounds on a government's use of individuals' personal data. Restricting interdepartmental access to such data will have an impact on the scope of AI's potential value creation. However, such regulation is necessary in order to enhance the uptake of AI without creating controversy between society and government. For example, although counterintelligence needs have never been more urgent, democratic stability is also dependent on the government's ability to safeguard the fundamental rights of its citizens to certain forms of privacy. Whilst there may never be a perfect balance, some prudence can be achieved: for example, the US model gives judges, not politicians, the responsibility for approval of privacy intrusion.

Enhance legal framework: A systematic legal framework on data collection and data use is a pressing need. Having a clear legal framework will enhance citizens' participation in data collection processes and enable citizens to monitor the transparency of data use.[178]

Write a data code of ethics and compliance: Adopt a code of ethics in order to ensure governments are following ethical standards and accountability controls. So what does an ethical AI look like? Bostrom and Yudkowsky (2014) suggest three qualities of an 'ethical algorithm': transparent to inspection, predictable to those they govern, and robust against manipulation.[179]

Transparency: As has been explained, an algorithm's transparency to inspection depends on its level of complexity. However, creative inspection techniques have been proposed that attempt to audit the decisions (outputs) made by the algorithm rather than the algorithm itself, through a trial and error method. For example, AI designed to decide who receives welfare benefits could be tested by making several applications with different combinations of characteristics with the aim of finding out what set of features result in the algorithm refusing or accepting an application ("crack the code"). This technique, also called algorithmic audits, was famously tried by Latanya Sweeney, who found statistically significant discrimination in Google's ad delivery algorithm.[180]

[178] Bertot, J. et al. (2010); Janssen, K. (2011)
[179] Bostrom, N., and Yudkowsky, E. (2014)
[180] Sweeney, L. (2013)

Predictability: Whether an algorithm is predictable is dependent on an ability to test them. The idea behind predictability is that, similar to the rule of law and the principle of *stare decisis* (i.e. that whenever possible judges must base their decisions on precedent), those governed by this principle can broadly predict what the outcome of their case is going to be.[181] With algorithms that make decisions for us or about us, the same principle should apply, and just as "law and ethics are normally domain specific; algorithmic accountability will have to be as well".[182]

Robust against manipulation: The final quality that an ethical AI must possess is to be robust against manipulation. This becomes increasingly important as use of AI increases: the more we delegate responsibilities to algorithms, the more vulnerable we are to their performance.

We accordingly propose two different components that must be included in the code of ethics:

a) **Justify outcomes before engaging in further data collection and analytics initiatives**. The added value to stakeholders must first be thoroughly studied before the commencement of data collection or an analytics programme.

b) **Produce an annual compliance report**. For example, the National Security Administration (NSA) has taken this initiative very seriously by creating the Civil Liberties and Privacy Office. The responsibility of the office is to conduct internal audits of the NSA operations and to ensure that each activity meets the Fair Information Practice Principles.[183] The published report should also include a clear selection method for which datasets are to be kept confidential or and which kept open to the public, with a justification on the added value generated for stakeholders by engaging in a particular activity. In doing so, governments will be able to generate greater public understanding around their technology initiatives and mitigate potential paranoia around data analytic activities.[184]

181 Bostrom, N., and Yudkowsky, E. (2014)
182 Rieder, B. (2015)
183 The fair information practice principles are internationally accepted for assessing the respect for privacy and are approved by the Privacy and Civil Liberties Oversight Board (Richards, 2014)
184 Mulgan, G. (2016)

VIII. References

Ackerman, Spencer (2016) Bill would expand FBI's warrantless access to online records, senators warn *The Guardian* Retrieved 28 June 2017 from https://www.theguardian.com/us-news/2016/may/26/fbi-intelligence-senate-bill-warrants-data-access

Adler, Tim (2014) Measure by measure: Algorithmic regulation is an idea whose time has come *Business Technology* Retrieved 29 June 2017 from http://business-technology.co.uk/2014/09/measure-by-measure-algorithmic-regulation-is-an-idea-whose-time-has-come/

Akerkar, Rajendra, ed. (2013) *Big Data Computing* Chapman & Hall

Akkiraju, R. (2015) IBM Watson Tone Analyzer – new service now available *IBM Watson* Retrieved 29 June 2017 from https://developer.ibm.com/watson/blog/2015/07/16/ibm-watson-tone-analyzer-service-experimental-release-announcement/

Altonji, Joseph G. and Rebecca M. Blank (1999) Race and gender in the labor market (in) *Handbook of Labor Economics*, Volume 3, Part C, 1999, pp. 3143-3259, Elsevier

"Artificial intelligence: opportunities and implications for the future of decision making" (2015) *UK Government Office for Science* Retrieved 29 June 2017 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf

"Artificial Intelligence, Robotics, Privacy and Data Protection" (2016) *European Data Protection Supervisor: 38th International Conference of Data Protection and Privacy Commissioners* Retrieved 29 June 2017 from https://edps.europa.eu/sites/edp/files/publication/16-10-19_marrakesh_ai_paper_en.pdf

Barry-Jester, Anna Maria et al (2015) Should Prison Sentences Be Based on Crimes That

Haven't Been Committed Yet? *FiveThirtyEight Politics* Retrieved 28 June 2017 http://fivethirtyeight.com/features/prison-reform-risk-assessment/ [http://perma.cc/Z6FK-H9UA].

Berg, Nate (2014) Predicting crime, LAPD-style *The Guardian* Retrieved 26 June 2017 from https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report

Berk, R. and J. Hyatt (2015) Machine learning forecasts of risk to inform sentencing decisions *Federal Sentencing Reporter*, 27(4) 222–228

Bershidsky, L. (2017) Machines Can Replace Millions of Bureaucrats *Bloomberg View*. Retrieved 29 June 2017 from https://www.bloomberg.com/view/articles/2017-02-09/machines-can-replace-millions-of-bureaucrats

Bertot, J. C., P. T. Jaeger and J. M. Grimes (2010) Using ICTs to create a culture of transparency? E-government and social media as openness and anti-corruption tools for societies *Government Information Quarterly* 27, 264–271

Bertot, J. C., P. T. Jaeger and D. Hansen (2012) The impact of polices on government social media usage: Issues, challenges, and recommendations *Government Information Quarterly* 29, 30–40.

Bertot, J. C., P. T. Jaeger, S. Munson and T. Glaisyer (2010) Engaging the public in open government: The policy and government application of social media technology for government transparency *IEEE Computer*, 43(11), 53–59

Betts, Peter and Charles Lound (2010) The application of alternative modes of data collection on UK Government social surveys *Office for National Statistics* Retrieved 29 June 2017 as PDF file from http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/data-collection-methodology/reports-and-publications/alternative-modes-of-data-collection/a-report-for-the-government-statistical-service.pdf

Birmingham Children's Hospital (2015) Quality Account 2014-15 *Birmingham Children's Hospital NHS Foundation Trust* Retrieved 29 June 2017 from http://www.bch.nhs.uk/sites/bch/files/document_files/bchft_quality_account_2014-15_final.pdf

Bostrom, N. and Yudkowsky, E. (2014) The ethics of artificial intelligence *in The Cambridge Handbook of Artificial Intelligence* (pp. 316-334) Cambridge University Press

Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies* Oxford University Press

Boyd, D., K. Levy, and A. Marwick (2014) The Networked Nature of Algorithmic Discrimination *Open Technology Institute* Retrieved 29 June 2017 from https://www.danah.org/papers/2014/DataDiscrimination.pdf

Brustein, J. (2016) Uber and Lyft Want to Replace Public Buses *Bloomberg Technology* Retrieved 29 June 2017 from https://www.bloomberg.com/news/articles/2016-08-15/uber-and-lyft-want-to-replace-public-buses

Bulger, Monica, Greg Taylor and Ralph Schroeder (2014) Data-Driven Business Models: Challenges and Opportunities of Big Data *Oxford Internet Institute* Retrieved 29 June 2017 from https://www.slideshare.net/DarylMcNutt1/businessmodelsforbigdata2014oxford

Burgess, Matt (2017) Despite the controversy, DeepMind signs up another NHS trust to use its Streams app *WIRED* Retrieved 27 June 2017 from http://www.wired.co.uk/article/deepmind-nhs-streams-deal

Cabinet Office (2015) Open Standards principles *GOV.UK*

Cairney, P. (2013) Policy Concepts in 1000 Words: The Policy Cycle and its Stages *Paul Cairney: Politics & Public Policy* Retrieved 29 June 2017 from https://paulcairney.wordpress.com/2013/11/11/policy-concepts-in-1000-words-the-policy-cycle-and-its-stages/

Campbell, W. K. and C. Sedikides (1999) Self-threat magnifies the self-serving bias: A meta-analytic integration *Review of General Psychology 3*(1) 23-43

Candler, G. (2014) Public Policy Theory *University of North Florida* Retrieved 29 June 2017 from http://www.unf.edu/~g.candler/PAD5384/03.pdf

Chalmers, John (1959) The Role of Government *Challenge* 8(3)27-30

Chang, Ai-Mei and P. K. Kannan (2008) Leveraging Web 2.0 in Government, *The IBM Center for the Business of Government* Retrieved 27 June 2017 from http://www.businessofgovernment.org/report/leveraging-web-20-government

Chapman, A. et al. (2016). Strengthening the Energy Policy Making Process and Sustainability Outcomes in the OECD through Policy Design. *Adm. Sci.* 6(9)

Charniak, Eugene and Drew McDermott (1985) *Introduction to Artificial Intelligence* Addison-Wesley

Collins, K. (2015) Government pays companies to monitor you on social media. *WIRED.* Retrieved 29 June 2017 from http://www.wired.co.uk/article/government-pays-companies-to-monitor-social-media-use

Copeland, M. (2016) What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning? *Nvidia* Retrieved 29 June 2017 from https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

Daly, E. and S. Singham (2012) Delivery 2.0: The new challenge for governments, *McKinsey and Company*

Danziger, Shai, Jonathan Levav and and Liora Avnaim-Pesso (2011) Extraneous factors in judicial decisions *Proceedings of the National Academy of Sciences* 108(17) 6889–6892 Retrieved 27 June 2017 from http://www.pnas.org/content/108/17/6889

"Data Science for Social Good" (2015) *The University of Chicago* Retrieved 29 June 2017 from https://dssg.uchicago.edu/projects/

Davenport, T. and Patil, D. (2012) Data Scientist: The Sexiest Job of the 21st Century *Harvard Business Review* Retrieved 27 June 2017 from https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

Davies, Will (2016) Robot Amelia – a glimpse of the future for local government *The Guardian* Retrieved 29 June 2017 from https://www.theguardian.com/public-leaders-network/2016/jul/04/robot-amelia-future-local-government-enfield-council

Desouza, Kevin C. and Jacob Benoy (2014) Big Data in the Public Sector: Lessons for Practitioners and Scholars *Administration & Society* Retrieved 27 June 2017 from http://journals.sagepub.com/doi/abs/10.1177/0095399714555751

Dewan, Shaila (2015) Judges Replacing Conjecture With Formula for Bail *The New York Times* Retrieved 27 June 2017 from https://www.nytimes.com/2015/06/27/us/turning-the-granting-of-bail-into-a-science.html

Digital Health (2017) Big Read: What does Google DeepMind want with the NHS? *Digital Health* Retrieved 27 June 2017 from https://www.digitalhealth.net/2017/03/deepmind-mustafa-suleyman-interview/

"Digital skills crisis" (2016). *House of Commons Science and Technology Committee* Retrieved 29 June 2017 from https://www.publications.parliament.uk/pa/cm201617/cmselect/cmsctech/270/270.pdf

Dorrier, Jason (2015) Exponential Medicine: Deep Learning AI Better Than Your Doctor at Finding Cancer *SingularityHub* Retrieved 29 June 2017 from https://singularityhub.com/2015/11/11/exponential-medicine-deep-learning-ai-better-than-your-doctor-at-finding-cancer/

Executive Office of the President (2014) Big Data and Privacy: A Technological Perspective *President's Council of Advisors on Science and Technology* Retrieved 28 June 2017 from https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf

Friedler, Sorelle. A., Scheidegger, Carlos, and Venkatasubramanian, Suresh (2016) On the (im) possibility of fairness *Cornell University* Retrieved 29 June 2017 from https://arxiv.org/abs/1609.07236

Gakrelidz, N. (2017) Predicting London Crime Rates Using Machine Learning *Dataiku* Retrieved 29 June 2017 from https://blog.dataiku.com/predicting-london-crime-rates-using-machine-learning

Gleason, D. (1990) Artificial Intelligence for Logistics Planning Systems *The Small Business Innovation Research (SBIR) program - U.S. Government* Retrieved 28 June 2017 from https://www.sbir.gov/sbirsearch/detail/214416

Glowacki, M. (2016) Nudging Cities: Innovating with Behavioral Science *Harvard Kennedy School ASH Center* Retrieved 28 June 2017 from http://datasmart.ash.harvard.edu/news/article/nudging-cities-innovating-with-behavioral-science-833

Goldin, Claudia and Cecilia Rouse (2000) Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians *The American Economic Review* 90(4) 715-741

Goldsmith, Stephen and Susan Crawford (2014) *The Responsive City: Engaging Communities Through Data-Smart Governance* Jossey-Bass

Griffin, Conor et al (2016) Advanced science and the future of Government *The Economist Intelligence Unit* Retrieved 29 June 2017 from https://www.cs.cityu.edu.hk/~hwchun/doc/news/Advanced_Science_future_of_gov_en.pdf

Guo, G. and X. Wang (2012) A study on human age estimation under facial expression changes *Proc. IEEE CVPR*, pp 2547–2553

Guthrie, J. (1982) Script to Print *Journal of Reading* 25(7), p. 717

Hanseth, O. and K. Lyytinen (2010) Design theory for dynamic complexity in information infrastructures: the case of building internet *Journal of Information Technology* 25, 1–19, Retrieved 28 June 2017 from http://heim.ifi.uio.no/~oleha/Publications/ISRinfrastructurefinal05-12-05.pdf

Hastie, Trevor, Robert Tibshirani and Jerome Friedman (2017) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (Second Edition)* Springer

Helbing, D. et al. (2017) Will Democracy Survive Big Data and Artificial Intelligence? *Scientific American* Retrieved 29 June 2017 from https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/

Holdren, J.P., et al (2016) Preparing for the Future of Artificial Intelligence *White House Executive Office of the President* Retrieved 29 June 2017 from https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

Howlett, Michael, M. Ramesh and Anthony Perl (2009) *Studying Public Policy: Policy Cycles and Policy Subsystems (Third Edition)* Oxford University Press

Howlett, Michael (2016) Matching Policy Tools & Their Targets: Beyond Nudges and Utility Maximization in Policy Design *Presentation at the European Consortium for Political Research Joint Session of Workshops* Retrieved 28 June 2017 from https://ecpr.eu/Filestore/PaperProposal/7aa0ca4c-c689-4f0f-bad1-12286042cd0c.pdf

IBM Newsroom (2014) IBM Watson Ushers in a New Era of Data-Driven Discoveries *IBM* Retrieved 28 June 2017 from http://www-03.ibm.com/press/us/en/pressrelease/44697.wss

IBM Newsroom (2017) Watson-Powered Tiatros Program for Veterans Achieves up to 73%

PTSD Program Completion *IBM* Retrieved 28 June 2017 from https://www-03.ibm.com/press/us/en/pressrelease/52251.wss

Jann, W. and K. Wegrich (2007) Theories of the Policy Cycle. In Fischer, F. et al (Eds.) *Handbook of Public Policy Analysis* Taylor and Francis Group, CRC Press

Janssen, K. (2011) The Influence of the PSI Directive on Open Government Data: an Overview of Recent Developments *Government Information Quarterly* 28, pp. 446-456

Jordan, M. I. and T. M. Mitchell (2015) Machine learning: Trends, perspectives, and prospects *Science* Vol 349: 6245 Retrieved 28 June 2017 from https://www.cs.cmu.edu/~tom/pubs/Science-ML-2015.pdf

Kahneman, Daniel, Andrew M. Rosenfield, Linnea Gandhi and Tom Blaser (2016) Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making *Harvard Business Review* Retrieved 28 June 2017 from https://hbr.org/2016/10/noise

Karsten, Jack and Darrell M. West (2016) Streamlining government services with bots *Brookings Institution* Retrieved 29 June 2017 from https://www.brookings.edu/blog/techtank/2016/06/07/streamlining-government-services-with-bots/

Kertcher, Zack and Ainat N. Margalit (2006) Challenges to Authority, Burdens of Legitimization: the Printing Press and the Internet *Yale Journal of Law and Technology*, 8(1) 17

Kleinberg, J., J. Ludwig, S. Mullainathan and Z. Obermeyer (2015) Prediction Policy Problems *American Economic Review* Papers and Proceedings (105)5

Klontz, Joshua C. and Anil K. Jain (2013) A Case Study on Unconstrained Facial Recognition Using the Boston Marathon Bombings Suspects *Michigan State University* Retrieved 28 June 2017 from https://pdfs.semanticscholar.org/01dc/2b7105e5f6420123e419839652c6886e3fa7.pdf

Kneese, Tamara, Alex Rosenblatand Danah Boyd (2014) Understanding Fair Labor Practices in a Networked Age *Data & Society* Retrieved 29 June 2017 from https://www.datasociety.net/pubs/fow/FairLabor.pdf

Knight, W. (2017) The Dark Secret at the Heart of AI *MIT Technology Review* Retrieved 28 June 2017 from https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/

Knight, W. (2017) A New Direction for Artificial Intelligence? *MIT Technology Review* Retrieved 28 June 2017 from https://www.technologyreview.com/s/603916/a-new-direction-for-artificial-intelligence/

Koerner, B. (2016) Inside the Cyberattack that Shocked the US Government *WIRED* Retrieved 29 June 2017 from https://www.wired.com/2016/10/inside-cyberattack-shocked-us-government/

Krizhevsky, Alex, Ilya Sutskever and Geoffrey E. Hinton (2012) ImageNet Classification with Deep Convolutional Neural Networks *Advances in Neural Information Processing Systems* Retrieved 28 June 2017 from https://www.nvidia.cn/content/tesla/pdf/machine-learning/imagenet-classification-with-deep-convolutional-nn.pdf

Labi, Nadya (2012) Misfortune Teller *The Atlantic* Retrieved 29 June 2017 from https://www.theatlantic.com/magazine/archive/2012/01/misfortune-teller/308846/

Latham, Erin (2016) Is the Government Hoarding Too Much Data? *Government Technology* Retrieved 28 June 2017 from http://www.govtech.com/opinion/Is-the-Government-Hoarding-Too-Much-Data.html

Levi, M, Sacks, A, and Tyler, T (2009) Conceptualizing Legitimacy, Measuring Legitimating Beliefs *American Behavioral Scientist* 53(3) 357

Li, C. and Yu, C. (2013) Performance Evaluation of Public Non-Profit Hospitals Using a BP Artificial Neural Network: The Case of Hubei Province in China *International Journal of*

Environmental Research and Public Health 10: 3619-33

Li, Xiaobai, Xiaopeng Hong, Antti Moilanen, Xiaohua Huang, Tomas Pfister, Guoying Zhao and Matti Pietikäinen (2015) Towards Reading Hidden Emotions: A comparative Study of Spontaneous Micro-expression Spotting and Recognition Methods *Cornell University* Retrieved 28 June 2017 from https://arxiv.org/abs/1511.00423

Lipton, Eric, Ian Austen and Sharon La Franiere (2013) Tension and Flaws Before Health Website Crash *The New York Times* Retrieved 28 June 2017 from http://www.nytimes.com/2013/11/23/us/politics/tension-and-woes-before-health-website-crash.html?pagewanted=all

Magdalena, R. et al (2015) Forecasting public expenditure by using feed-forward neural networks *Economic Research-Ekonomska Istraživanja*, *28*(1) 668-686.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh and Angela Hung Byers (2011) Big data: The next frontier for innovation, competition, and productivity *McKinsey Global Institute* Retrieved 28 June 2017 from http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation

McCormick, J. (2016) Predictions 2017: Artificial Intelligence Will Drive The Insights Revolution *Forrester Research* Retrieved 29 June 2017 from https://go.forrester.com/wp-content/uploads/Forrester_Predictions_2017_-Artificial_Intelligence_Will_Drive_The_Insights_Revolution.pdf

McCullough, Aoife (2015) The Legitimacy of States and Armed Non-State Actors *University of Birmingham* Retrieved 28 June 2017 from http://www.gsdrc.org/wp-content/uploads/2015/08/Legitimacy.pdf

Ministry of Justice (2015) Proven Re-offending Statistics Quarterly Bulletin: January to December 2013, England and Wales *Ministry of Justice Statistics Bulletin* Retrieved 28 June 2017 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/472524/proven-reoffending-2013.pdf

Mitchell, Caroline, Paul Meredith, Matthew Richardson, Peter Greengross and Gary B. Smith (2015) Reducing the number and impact of outbreaks of nosocomial viral gastroenteritis: time series analysis of a multidimensional quality improvement initiative *Portsmouth Hospitals NHS Trust* Retrieved 28 June 2017 from http://eprints.bournemouth.ac.uk/24304/1/BMJ%20QS%20norovirus%20version%20for%20BU%20BRIAN%20website%20%20BURO.pdf

Mohler G. O., M. B. Short, Sean Malinowski, Mark Johnson, G. E. Tita and Andrea L. Bertozzi Randomized Controlled Field Trials of Predictive Policing *Journal of the American Statistical Association* 110 Issue 512, 1399-1411 Retrieved 28 June 2017 from http://www.tandfonline.com/doi/abs/10.1080/01621459.2015.1077710

"More accountability for big-data algorithms" (2016) *Nature 537*: 449

Moreno, Hugo (2014) The Power of Data Analytics to Transform Government *Forbes* Retrieved 28 June 2017 from https://www.forbes.com/sites/forbesinsights/2014/10/01/the-power-of-data-analytics-to-transform-government/#6adb265466a1

Mulgan, Geoff (2016) A machine intelligence commission for the UK: how to grow informed public trust and maximise the positive impact of smart machines *Nesta* (delivered to) *The Alan Turing Institute* Retrieved 29 June 2017 from https://www.nesta.org.uk/sites/default/files/a_machine_intelligence_commission_for_the_uk_-_geoff_mulgan.pdf

"National and Local Government" (n.d.) *Rother District Council* Retrieved 29 June 2017 from http://www.rother.gov.uk/article/2689/National-and-local-government

National Institute of Justice (2014) Recidivism, *Office of Justice Programs* Retrieved 28 June 2017 from https://www.nij.gov/topics/corrections/recidivism/Pages/welcome.aspx

Neill, Daniel B. (2012) New directions in artificial intelligence for public health surveillance *IEEE Intelligent Systems* 27(1) 56-59

Neill, Daniel B. (2013) Using artificial intelligence to improve hospital inpatient care *IEEE Intelligent Systems* 28(2) 92-95

"New Media Monitoring and Evaluation agreement" (2015) *Gov.UK* Retrieved 29 June 2017 from https://www.gov.uk/government/news/new-media-monitoring-and-evaluation-agreement

Nowozin, S. (2017) Personal interview with author by Adira E. Levine *Centre for Public Impact*

"Of Prediction and Policy" (2016) *The Economist* Retrieved 29 June 2017 from http://www.economist.com/news/finance-and-economics/21705329-governments-have-much-gain-applying-algorithms-public-policy

Önder, E. et al (2013) Forecasting macroeconomic variables using artificial neural network and traditional smoothing techniques *Journal of Applied Finance & Banking 3*(4) 73-104

Painter, Martin et al (2010) *Tradition and Public Administration* Palgrave Macmillan

Pearlstein, Steven (2014) The federal outsourcing boom and why it's failing Americans *The Washington Post* Retrieved 28 June 2017 from https://www.washingtonpost.com/business/the-federal-outsourcing-boom-and-why-its-failing-americans/2014/01/31/21d03c40-8914-11e3-833c-33098f9e5267_story.html?utm_term=.7140c3082047

Pierre, Jon, Asbjørn Røiseland and Annelin Gustavsen (2011) Legitimacy by performance in democratic systems? Preparing for empirical analysis *Nordisk Kommuneforskerkonferanse*

"Policy Making and Policy Implementation" (n.d.). *Texas Policy Project, University of*

Texas. Retrieved 29 June 2017 from
https://texaspolitics.utexas.edu/archive/html/bur/features/0303_01/policy.html

Poole, D. L. and A.K. Mackworth (2010) *Artificial Intelligence: Foundations of Computational Agents* Cambridge University Press

Powles, Julia and Hal Hodson (2017) DeepMind and healthcare in an age of algorithms *Health and Technology* Retrieved 27 June 2017 from
https://link.springer.com/article/10.1007/s12553-017-0179-1

Rahmes, Mark, Mike McGonagle, J. Harlan Yates, Ronda Henning and Jay Hackett (2013) An efficient visualization method for analyzing biometric data *Biometric and Surveillance Technology for Human and Activity Identification* Retrieved 28 June 2017 from
http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1693605

Rich, E. and Knight, K. (1991) *Artificial Intelligence* McGraw-Hill

Richards, Rebecca (2014) Prepared Remarks: Privacy and Civil Liberties Oversight Board Forum *National Security Agency* Retrieved 28 June 2017 from
https://www.nsa.gov/about/civil-liberties/resources/assets/files/PCLOB_Remarks_20141112.pdf

Rieder, Bernhard (2015) On the Diversity of the Accountability Problem: Machine Learning and Knowing Capitalism *Universiteit van Amsterdam Mediastudies Department* Retrieved 28 June 2017 from http://www.law.nyu.edu/sites/default/files/rieder_NY_feb2015.pdf

Ritter, Nancy (2013) Predicting Recidivism Risk: New Tool in Philadelphia Shows Great Promise *NIJ Journal* Issue No. 271 Retrieved 28 June 2017 from
https://www.ncjrs.gov/pdffiles1/nij/240696.pdf

Rogers, Simon (2012) Public v private sector pay: who earns more? *The Guardian* Retrieved 28 June 2017 from https://www.theguardian.com/news/datablog/2012/mar/27/public-private-sector-pay

Rothstein, Bo (2009) Creating Political Legitimacy: Electoral Democracy Versus Quality of Government *American Behavioral Scientist* 53(3) 3

Rubin, J., (2012) The Printing Press, Reformation, and Legitimization *Chapman University*, p. 16

Rudin, C. (n.d.) Predictive Policing Using Machine Learning to Detect Patterns of Crime *WIRED* Retrieved 29 June 2017 from https://www.wired.com/insights/2013/08/predictive-policing-using-machine-learning-to-detect-patterns-of-crime/

Russell, Stuart and Peter Norvig (2013) *Artificial Intelligence: A Modern Approach (Third Edition)* Prentice Hall

Santelices, M. V., and M. Wilson (2010). Unfair treatment? The case of Freedle, the SAT, and the standardization approach to differential item functioning *Harvard Educational Review* 80(1) 106–134

Saria, Suchi, Anand K. Rajani, Jeffrey Gould, Daphne Koller and Anna A. Penn (2010) Integration of Early Physiological Responses Predicts Later Illness Severity in Preterm Infants *Science Translational Medicine* 2(48) 48-65

SAS (2013) Big Data Analytics skills development in the UK is vital to realise business benefits and economic impact *SAS* Retrieved 29 June 2017 from https://www.sas.com/en_gb/news/press-releases/2013/november/e-skills-Big-Data-Analytics-skills-development-in-the-UK-is-vital-to-realise-business-benefits-and-economic-impact.html

Savard, J. F. and R. Banville (2012) Policy Cycles *Encyclopedic Dictionary of Public Administration* Retrieved 29 June 2017 from http://www.dictionnaire.enap.ca/dictionnaire/docs/definitions/definitions_anglais/policy_cycles.pdf

Schofield J. and J. Fershau (2007) Committees Inside Canadian Legislatures in Dobuzinskis,

L. et al. (Eds.) *Policy Analysis in Canada* University of Toronto Press

Schrater, P. (n.d.) Reinforcement Learning *University of Minnesota* Retrieved 28 June 2017 from http://gandalf.psych.umn.edu/users/schrater/schrater_lab/courses/AI2/rl1.pdf

Scioli, F. P. (1979) Problems and prospects for policy evaluation *Public Administration Review* 39(1) 41-45

Shead, Sam (2017) Why Google Deepmind's work with the NHS is being investigated by the regulators *Business Insider UK* Retrieved 28 June 2017 from http://uk.businessinsider.com/why-google-deepminds-work-with-the-nhs-is-being-investigated-by-the-regulators-2017-5?r=UK&IR=T

Sommerfeld, Chris (2015) The Rise of the Data Scientists *Data-Smart City Solutions* Retrieved 28 June 2017 from http://datasmart.ash.harvard.edu/news/article/the-rise-of-the-data-scientists-611

Spangler, Scott et al (2014) Automated hypothesis generation based on mining scientific literature *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp 1877-1886

Sparrow, Andrew and Hélène Mulholland (2011) Cameron accused of putting NHS on sale over plans for life sciences *The Guardian* Retrieved 28 June 2017 from https://www.theguardian.com/society/2011/dec/05/cameron-nhs-sale-life-sciences

Steele, Peggy (2015) 6 strategies for closing the big data talent gap *World Economic Forum* Retrieved 28 June 2017 from https://www.weforum.org/agenda/2015/10/6-strategies-for-closing-the-big-data-talent-gap/

Stein, Amanda (2014) IBM Watson Joins MD Anderson in Cancer Fight *Texas Medical Center* Retrieved 28 June 2017 from http://www.tmc.edu/news/2014/08/ibm-watson-joins-md-anderson-in-cancer-fight/

Stern, S. (2013) Connecting Millions of Americans with Health Coverage: The 2013-2014 Opportunity *Enroll America* Retrieved 29 June 2017 from http://www.nationalhealthcouncil.org/sites/default/files/NHC_Files/Pdf_Files/Stern.pdf

Subroto, A. (2011) Understanding Complexities in Public Policy Making Process through Policy Cycle Model: A System Dynamics Approach *System Dynamics Society* Retrieved 29 June 2017 from https://www.systemdynamics.org/conferences/2012/proceed/papers/P1067.pdf

Susskind, R. and D. (2015) *The Future of the Professions: How Technology Will Transform the Work of Human Experts* Oxford University Press

Sweeney, L. (2013) Discrimination in online Ad delivery *Communications of the ACM* 56(5), 44–54 Retrieved 27 June 2017 from https://arxiv.org/ftp/arxiv/papers/1301/1301.6822.pdf

Thwaites, Emma (2012) Prescription Savings Worth Millions Identified by ODI incubated company *Open Data Institute* Retrieved 28 June 2017 from https://theodi.org/news/prescription-savings-worth-millions-identified-odi-incubated-company

Tversky, A. and D. Kahneman (1974) Judgment under uncertainty: Heuristics and biases *Science* 185, 1124-1131

University of Oxford (2014) Giant £132m investment in cancer research *Oxford University* Retrieved 28 June 2017 from http://www.ox.ac.uk/news/2014-10-23-giant-%C2%A3132m-investment-cancer-research-oxford-university

Urban, T. (2015) The AI Revolution: The Road to Superintelligence *Wait But Why* Retrieved 29 June 2017 from http://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html

Varian, Hal R. (2014) Big Data: New Tricks for Econometrics *Journal of Economic Perspectives* 28(2) 3-28

Versluis, E. et al. (2011) *Analyzing the European Union Policy Process* Palgrave Macmillan

Waddell, K. (2015) The NSA's Bulk Collection Is Over, but Google and Facebook Are Still in the Data Business *The Atlantic* Retrieved 28 June 2017 from https://www.theatlantic.com/politics/archive/2015/06/the-nsas-bulk-collection-is-over-but-google-and-facebook-are-still-in-the-data-business/458496/

Waddell, K. (2016) Half of American Adults Are in Police Facial-Recognition Databases *The Atlantic* Retrieved 29 June 2017 from https://www.theatlantic.com/technology/archive/2016/10/half-of-american-adults-are-in-police-facial-recognition-databases/504560/

Warner, M. E. and J.E. Pratt (2005) Spatial diversity in local government revenue effort under decentralization: a neural-network approach *Environment and Planning C: Government and Policy* 23(5) 657-677

West, J. (2016) Microsoft's disastrous Tay experiment shows the hidden dangers of AI *Quartz* Retrieved 29 June 2017 from https://qz.com/653084/microsofts-disastrous-tay-experiment-shows-the-hidden-dangers-of-ai/

Wilson, J. Q. (1989) *Bureaucracy: What government agencies do and why they do it* Basic Books

Wong, J, I. (2017) DeepMind's access to UK health data shows how tech could outgun privacy laws *Quartz* Retrieved 29 June 2017 from https://qz.com/934137/googles-goog-deepmind-got-too-much-health-data-from-britains-nhs-research-paper-says/

Yang, Y. et al. (2016) An improved grey neural network forecasting method based on genetic algorithm for oil consumption of China *Journal of Renewable and Sustainable Energy* 8(2)

Yiu, Chris (2012) The Big Data Opportunity: Making government faster,smarter and more personal *Policy Exchange*

Zeng, Jiaming, Berk Ustun and Cynthia Rudin (2016) Interpretable Classification Models for

Recidivism Prediction *Journal of the Royal Statistical Society* Retrieved 28 June 2017 from https://arxiv.org/pdf/1503.07810.pdf