OECD Working Papers on Public Governance No. 33

A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance

Charlotte van Ooijen,

Barbara Ubaldi,

Benjamin Welby

https://dx.doi.org/10.1787/09ab162c-en

OECD

Data The first was a po

Organisation for Economic Co-operation and Development

GOV/PGC/EGOV(2019)3

Unclassified **English - Or. English**

PUBLIC GOVERNANCE DIRECTORATE
PUBLIC GOVERNANCE COMMITTEE

OECD Digital Government Project

A data-driven public sector

Enabling the strategic use of data for productive, inclusive and trustworthy governance

OECD Working Papers on Public Governance

Disclaimer: OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to the Public Governance Directorate, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France

Ubaldi, Barbara-Chiara
Van Ooijen, Charlotte
Welby, Benjamin

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Foreword

It is well understood that data play an important role in the way in which twenty-first century public services are administered. In fact, the successful implementation of the digital government approaches which the OECD recommends rely on this foundation of 'data' to deliver on the potential of twenty-first century government.

Over the last decade there has been an increased awareness of the role of data through the Open Government Data movement. Its success in seeing governments open up its information for reuse both inside, and outside, the public sector has proven significant in putting data at the heart of the transformation agenda.

This Working Paper argues that governments now need to go further. The question of how data can be collected, processed, shared and reused is one that countries need to grapple with and place at the core of their thinking about digital transformation. Consequently, governments will move from a focus on the external publication of data towards a highly adept public service that is skilled in the recognition and use of data as a core component of a highly functioning state and the effective design and delivery of its activities.

The Public Governance Directorate has several teams focused on the Reform of the Public Sector. For all these teams the question of digital transformation is critical to the future capacity of governments to address the concerns of their citizens and none more so than the Digital Government Unit.

This analysis of the concept of a data-driven public sector draws on the OECD knowledge base and exchange of good practice, experience and strategies developed through the OECD Working Party of Senior Digital Government Officials (E-Leaders) and in particular the work of the Thematic Group on a data-driven public sector. The Working Paper builds upon the analytic framework provided by the OECD Recommendation of the Council on Digital Government Strategies (2014) and the experience of projects conducted over the last 15 years across OECD member countries and other countries.

The Working Paper *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance* was prepared by the OECD Directorate for Public Governance (GOV), under the leadership of Marcos Bonturi.

This document was produced by the OECD Reform of Public Sector Division (GOV/RPS), headed by Barbara-Chiara Ubaldi, Acting Head of Division. It benefitted from her strategic orientation and revisions in that capacity and as lead on Digital Government, the Data-driven Public Sector and Open Data work. This document was drafted by Charlotte van Ooijen and Benjamin Welby, Policy Analysts - Digital Government and Data driven public sector. Colleagues from within the OECD including Lucia Chauvet, Christian Reimsbach-Kounatze and David Gierten reviewed and provided comments. Raquel Páramo provided editorial and administrative support.

Table of contents

Foreword **2**

Executive Summary **4**

1. The context and imperative for a Data-Driven Public Sector **6**

2. Data-driven public value creation **11**

3. Opportunities **17**

4. Challenges **28**

5. Enabling frameworks **46**

6. Conclusion **53**

Figures

Figure 1. Data governance in the public sector 8
Figure 2. Government Data Value Cycle 11
Figure 3. Algorithmic data processing 13
Figure 4. Opportunities of a data-driven public sector 18
Figure 5. Challenges in fostering a data-driven public sector 29
Figure 6. Norway's base registries 33
Figure 7. Confidence in national government in 2015 and its change since 2007 38
Figure 8. Existence of a Chief Data Officer for central government in 2014 compared to 2016 48
Figure 9. Enabling frameworks for a data-driven public sector 52

Boxes

Box 1. The role of data analytics in the fight against fraud 24
Box 2. Innovative performance monitoring by the SAI in India 27
Box 3. Digital skills in the United Kingdom government 35
Box 4. The ICT Commissioning Playbook 37
Box 5. The US Federal Data Strategy 47
Box 6. The Italian Data & Analytics Framework (DAF) 50

Executive Summary

The digital transformation of society has increasingly changed the way that economies function, the social inclusion challenges facing governments, and the way in which our democracies operate. As governments explore their response to these issues they have understood the growing importance of the value of data.

Some of these efforts have focused on the importance of open government data, leading to the publication of public sector data to stimulate private sector innovation, provide opportunities for the economy at large and increase government accountability. In other cases, the re-use of data internally within government and for external impact on the experience of citizens is patchier. However, both scenarios do not yet deliver on the potential for governments to use data as a foundation for improved policy making, service delivery and ongoing performance management.

It is in this context that the OECD is exploring the concept of a data-driven public sector (DDPS), one which recognises data as an asset, integral to policy making, service delivery, organisational management and innovation. The strategic approach that governments take to building a DDPS can have a positive impact on the results they deliver by promoting evidence-led policy making and data-backed service design as well as embedding good governance values of integrity, openness and fairness in the policy cycle.

This working paper considers the power of data to transform our governments, including its use as a strategic asset to increase societal well-being and its potential democratising effect on public discourse. However, it is important to identify how governments can move on from pockets of good practice to whole-of-government approaches that are well supported internally and which also find favour with the public.

The paper discusses the opportunities created by a DDPS in three important areas. First, the ability of governments to use data to be better prepared for the future. This explores how data can be used to predict trends and patterns in order to mitigate emerging risks and respond to developing crises. The second area of opportunity is around the design and delivery of policy and services. Data provide important ways of understanding problems, engage the public and provide access to insights for improving public services that meet user needs, while creating the conditions for robust, evidence-based policy making. The third area of opportunity that the paper explores is performance management in terms of greater public sector productivity and better evaluation of policies and impact.

The paper subsequently identifies challenges governments may face in establishing a DDPS. The first of these concerns the availability, quality and relevance of data. The second is sharing data internally, which involves tackling organisational resistance, finding ways to communicate what is available internally and solving barriers to interoperability. The third set of challenges concerns the skills and capabilities that governments need to make the most of data. Finally, the critical question of legitimacy and public trust is considered, with a focus on ethical use of data by public sector organisations, privacy, transparency and the risks that governments and citizens need to be aware of when implementing a data-driven public sector.

The paper concludes with the discussion of the need for coherent strategic approaches that reflect the role of data across the entire public sector, not only from a policy point of view but from an operational and practical perspective.

Further work

This Working Paper sets out the framework to analyse existing practices, identify policy implications and efforts required by governments to maximise the benefits of the DDPS, and advance in policy options. However, this raises questions about how countries should go about implementing some of these ideas and tackling the barriers that they may come across.

As such, this Working Paper paves the way to a Report the Secretariat anticipates publishing in 2019 that considers the comparative experiences of OECD countries in areas such as:

- creating the conditions for a DDPS – starting with the establishment of the necessary model of governance to secure leadership, skills and technological capabilities;
- adopting ethical approaches to the management and use of government data (particularly citizen data) and the need to safeguard and explore emerging digital rights;
- developing business cases for DDPS by identifying potential gains, proposing mechanisms to achieve these gains and highlighting priority areas of investment;
- mapping the possible costs of inaction and inability to mainstream DDPS practices, for example in terms of financial loss (e.g. the inability to address data-related fraud) or exacerbating the trust deficit (e.g. the inability of responding to expectations of citizens).

1. The context and imperative for a Data-Driven Public Sector

This chapter provides the context for the OECD's work on the Data-Driven Public Sector. The first part introduces the ambitions, which people have often had for the role of data in the delivery and performance of government by discussing how data can be viewed as a strategic asset. Followed by the way in which data have been increasingly democratised by assuring its availability and accessibility to the vast majority of citizens, in order to maximize value creation.

The second part of the chapter discusses how governments can move from the hype of data to having impact. It suggests that whilst there have been pockets of good practice contributing to improved outcomes for citizens, it is now essential for countries to develop whole of government approaches that put the value of data at the heart. This paper hopes to provide the materials for governments to make that case whilst ensuring that citizens are protected and that the outcomes lead to greater trust in government from the public.

Finally, the chapter sets out some of the questions that will be explored, the structure for the Working Paper and the broader OECD context into which it fits.

1.1. The potential power of data

What if more poor families were to receive the state benefits they need? What if more patients' lives can be saved? What if more children could be protected from situations of abuse? What if more people can be rescued from imminent flooding? What if more terrorist attacks could be prevented?

What if government policies delivered better results and improved the wellbeing of all?

The realisation of better policy results depends in large part on taking the right decisions at the right time. Whether it is a doctor deciding on the course of treatment, a policy maker designing a social benefits package, a tax inspector deciding where to make his next house call, a risk manager deciding which rescue team to send where at what moment in time, an anti-terrorism coordinator deciding on preventive action, or a parent choosing which school their child should attend, these decisions all depend on receiving the right insights at the right time to make an informed choice.

Long-term trends like ageing populations, rising inequality and climate change alongside the unpredictable challenges of issues like migration and security affect citizens' well-being and resilience. This requires governments to develop long-term vision, make evidence-based investment and coordinate their action. Governments are increasingly aware of how the better use of data can support a response to these challenges, while also improving the productivity, performance and inclusiveness of policies and services.

A Data-Driven Public Sector (DDPS) transforms the design, delivery and monitoring of public policies and services through the management, sharing and use of data. Using data as a strategic asset is crucial for governments to boost public sector intelligence and, as a result, increase the capability of developing policies and services that are sustainable in the long-term and as inclusive and trusted as possible.

A DDPS starts with the recognition of data as a fundamental building block for those seeking to implement and enable digital government in their countries. Thanks to the increasing wealth of data made possible in the digital age and the widespread adoption of Information and Communication Technologies (ICTs), data and analytical tools have become widely available, providing a wealth of opportunity to improve decision-making. Besides textbooks, expert advice or routine procedures, decision makers can access tailor-made, context specific information and knowledge informed by an analysis of data.

Digital technologies and the data revolution change the economic and societal landscape. This affects a range of sectors such as research, education and industry, forcing governments to develop new approaches to challenges of skills, approaches to researching policy, and the next production revolution. Furthermore, it affects the functioning of the public sector itself, requiring a reassessment of public governance and a reflection on government's role in the information society.

1.2. The necessity to move from hype to impact

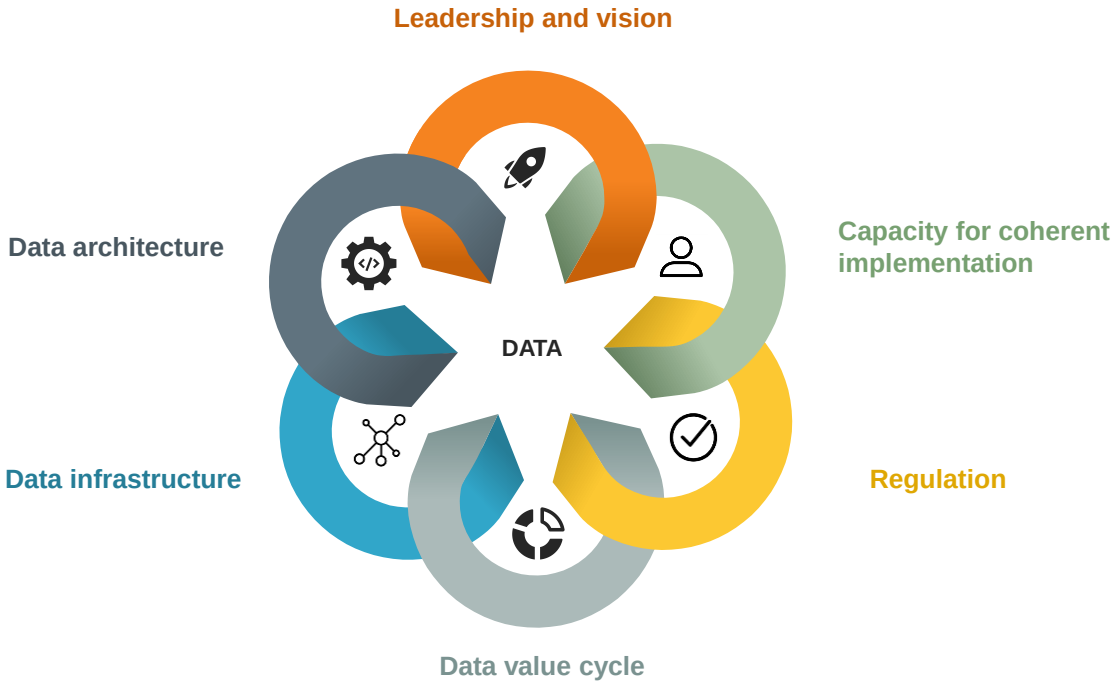
There is no shortage of compelling examples demonstrating how a smarter use of data enables innovation in public policies and services.

In France, the Department of Labour and the Secretariat-General for Government Modernisation opened access to underused government data related to unemployment, one of the most significant social issues in the country. This allowed non-profit organisation Bayes Impact to work with data from *Pôle emploi*, the French unemployment agency, and explore the use of algorithms to develop the citizen-led digital service called *Bob emploi* (https://www.bob-emploi.fr/).

The public sector can be an important contributor to the provision of health care whether through setting the direction of public health policy or direct provision. The leading cause of illness and death worldwide, cardiovascular disease might be better addressed if the experienced judgment of human experts in risk prediction were augmented by data-driven evidence. A 2017 study from the United Kingdom shows that machine-learning significantly improves the accuracy of cardiovascular risk prediction, increasing the number of patients identified who could benefit from preventive treatment, while avoiding unnecessary treatment of others (Weng et al., 2017[1]). Such an intervention ensures the targeting of delivery resources in a more effective and efficient way allowing other priorities to be addressed.

These two examples are indicative of the potential for data-driven approaches to transform lives. Nevertheless, it is important to retain a critical stance in terms of how governments use data to ensure that the impact is sustainable, inclusive and trusted. DDPS initiatives are being developed in many sectors but often this is happening in siloes, according to existing institutional logic. The importance of analytics in delivering better outcomes in one policy area or level of government is valuable but the increased benefits of extending that approach throughout the public sector demand system approaches and joint solutions.

Figure 1. Data governance in the public sector



Source: Digital Government Review of Argentina (OECD, 2019[2])

Data governance (Figure 1) is the exercise of authority, control, and shared decision making (planning, monitoring and enforcement) over the management of data assets (Ladley, 2012[3]) either within one organisation or across different organisations that share an interest for common data assets. A data governance model must support existing and new processes to ensure the proper management and protection of data, proper production and usage of data through its life cycle in a collaborative, federated approach to managing valuable data assets (Ghavami, 2015[4]).

Enabling data-driven activities require a recognition of the breadth of data governance to cover a strategic approach and effective management models that help address the common challenges to unlocking the value of data, such as training and skills. Additionally, data sharing is a critical enabler that necessitates the adoption of government wide approaches. It is essential for a whole-of-government approach to be applied when it comes to defining and implementing risk-based data governance, including the development of frameworks to support ethical and transparent approaches to handling and using data within the public sector, in particular where those are personally identifiable and sensitive.

Yet making the case for DDPS is not immediately apparent; in order to secure government investment the gains for the public sector must clearly support the case for governments to intervene. This paper aims to help governments develop the business case for DDPS by identifying its potential gains, proposing mechanisms to achieve these gains and highlighting priority areas of investment. Alongside those benefits, it will address the possible costs of inaction in terms of financial loss (for example, the inability to address data-savvy fraud) or exacerbating the trust deficit (for example, the inability of responding to expectations of citizens). Most governments lack both an overall vision and strategic approach to how they use public sector data as key enablers of broader public sector

transformation and modernisation agendas, and improve citizens' well-being in a sustainable, inclusive and trusted way.

Success stories provide compelling arguments for promoting DDPS. The argument is compelling because of the potential impact: several cases have shown that the exchange and analysis of data can directly enhance people's well-being or even save lives such as the efforts coordinated by the Humanitarian Data Exchange during the Ebola crisis of 2014 (Griliopoulos, 2014[5]). However, such arguments can also be dangerous. They can be used to legitimise data sharing or surveillance legislation on the basis of protecting the public whilst giving others reason to distrust government and fear risks to privacy such as the debate around the UK's Investigatory Powers Act of 2016 (Schafer, 2016[6]). These arguments can also lead to the assumption that data sharing and analytics automatically create public value, while bounded rationality in decision-making and institutional factors should be taken into account as well. Simply having the 'right' information at the 'right' time does not automatically lead to the 'right' decision because the decision-maker, whether human or machine, may lack the legal right, ability, ethics, conviction, motivation or political support to take the right decision. Moreover, the definition of the "right decision" may not only depend on the decision made but the process leading up to it: was the decision taken in a fair, democratic and inclusive way?

Early e-Government efforts focused on efficiency improvements through automated data processing (OECD, 2003[7]) with the idea that existing paper-based processes and service delivery methods, designed in the analogue age could gain in terms of transparency and efficiency when transferred on electronic platforms. However, today data management and analytics provide the opportunity to assess the efficiency of existing processes and methods and to rethink and change them accordingly. Technology and the digitisation of societies and governments are generating massive amounts of data. The increased availability, quantity, complexity and production rate of data, together with revolutionary developments in data analytics have expanded the possible applications of data well beyond existing electronic service delivery as understood years ago. Intelligent data usage offers a myriad of possibilities to fundamentally transform public sector activities, how services are designed, delivered and monitored, Data analytics can reveal social and economic trends and incite a transition from reactive to more proactive and forward-looking data-based policymaking and service delivery.

Integrated analysis of data concerning service users – both businesses and citizens – and contextual (big) data may facilitate a move from citizen-centred to more citizen-driven public service design and delivery. Better integration and analysis of up-to-date, varied and detailed performance data could boost public sector productivity by enabling a change in the focus of policy evaluations from one-time assessments in the moment to more continuous performance improvement. Nevertheless, data do not provide public value in isolation. The necessary conditions for data governance (see Figure 1) must be created where data can be managed in an intelligent way, enabling the necessary transition to a public sector capable of sharing and using data to design and implement more sustainable, inclusive and trustworthy policies and services (OECD, 2018[8]; Ubaldi, 2013[9]; OECD, 2017[10]; OECD, 2019[2]).

Data can be the fuel to power the digital transformation of the public sector, but if governments want the data revolution to have a substantial positive impact on people's well-being while contributing to public trust, it is important to ensure the flow of data to decision makers by simplifying access and removing any barriers that might exist. Whilst getting the right information to the right decision-makers at the right time has an impact, a

focus on the result of policy is not enough; the limitations of rationalistic approaches to decision-making can contribute to a loss of public trust. The OECD report *Trust and Public Policy* (OECD, 2017[11]) identifies two critical policy levers impacting citizens' trust in public institutions: policy competence, and embedding the values of good governance into policymaking. DDPS has the potential to provide a significant boost to efforts in both of these areas.

1.3. Structure and scope

This paper aims to present and discuss some of the main opportunities for governments to improve their capacities to manage and use data as a strategic asset as a means of positively impacting people's lives. Governments which do this pave the way to pursue sustainable, inclusive and trustworthy data-driven public value creation.

While the OECD work on Open Government Data (OECD, 2015[12]; OECD, 2016[13]; Ubaldi, 2013[9]) puts the spotlight on public institutions in the role of data producers, the primary focus of DDPS is on how public institutions, at the international (including across borders), national, regional and local levels of government, and within different policy sectors, can better fulfil their role as data users. The main question addressed is how recognising data as a strategic resource can transform the way in which governments detect policy issues, design and deliver policies and services that respond to those issues, and monitor and evaluate the performance of those responses, with the overall objective of increasing well-being for all. The involvement of other societal stakeholders will prove to be of great importance, but will be discussed from the perspective of data use in the public policy process.

In order to assist countries in assessing their respective data-driven landscapes and determining what steps to take towards maximum value creation while adequately addressing the risks, the following questions are addressed in this paper:

- Through what process can data support decision-making for policies, services and organisational management?
- What is the potential gain of data-driven decision-making? Is it possible to identify the opportunities?
- What are the challenges when implementing DDPS?
- What conditions are needed to support sound data-driven decision-making throughout the whole of government?

This Working Paper complements the analysis and discussion points raised in two other Working Papers, one on the State of the art on the use of emerging technologies in the public sector and the other on the impact of digital government on citizen well-being (Welby, 2019[14]; Ubaldi et al., 2019[15]).

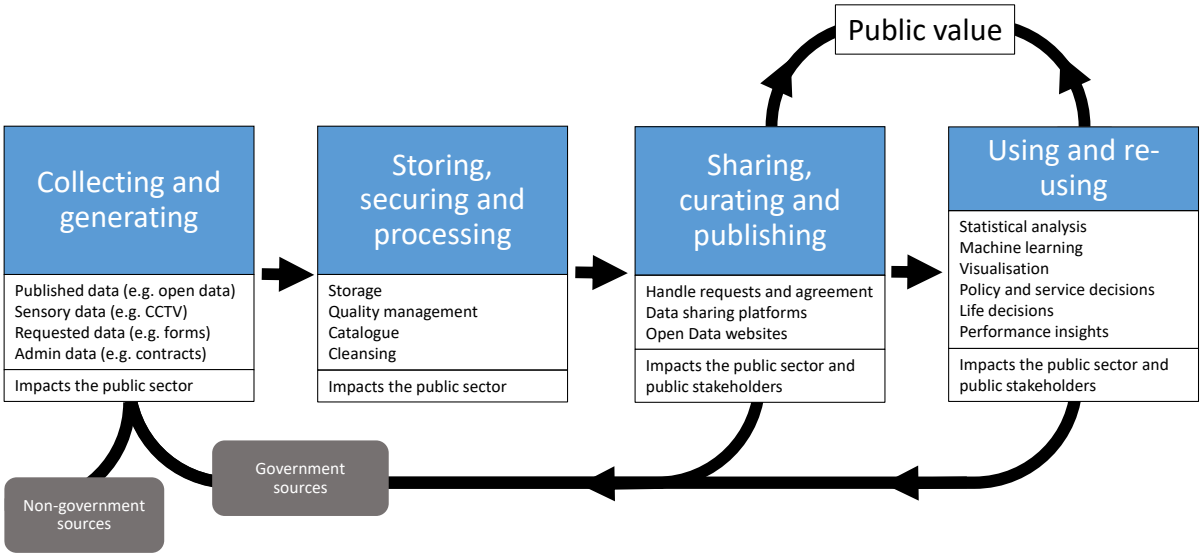
2. Data-driven public value creation

This chapter provides an overview of several components involved in creating public value through the use and application of data. This chapter highlights the importance to government of being clear about the role it expects data to play in decision making and the efforts that are needed to turn data into information, both initially and on an ongoing basis, to support that. The second half of this chapter explores in more detail the characteristics of data in terms of where it is sourced from, how it is used and those who are involved in its use and re-use.

2.1. Government Data Value Cycle: supporting public decision-making

The improved management of the government data value cycle (Figure 2) can help policy makers and public officials increase their effectiveness by enhancing their capability to gather insight on existing policy problems and different stakeholders; foresee new trends and needs; design and adapt innovative policy approaches; monitor the activities undertaken and policies implemented; and manage the resources (financial, time, human and material) mobilised to address policy challenges.

Figure 2. Government Data Value Cycle



This application of analytical methods which turn data (raw, isolated and unstructured datasets) into information (where relationships between data are identified) and knowledge (understanding those relationships) provides the basis for decision-making at the strategic, tactical and operational levels of government (taking action) (OECD, 2015[16]).

Nevertheless, boosting public sector intelligence and creating public value does not happen in a linear value chain but through a cycle, which involves feedback loops throughout the value creation process. Data can inform and affect the nature of decision-making processes

which in turn can lead to the production and collection of different or more data (OECD, 2015[16]).

For example, knowledge about the cause of a disease may cause decision-makers to focus data collection efforts on both the disease, and the spread of its causes. This allows for the development of preventative measures targeted at the cause and its spread, in conjunction with a continued monitoring of the disease, informing the effectiveness of these measures on both the cause, and the disease, through this additional data collection. As such, both the output, and process, of decision-making are strengthened through a DDPS approach.

2.1.1. What data?

What data are at the disposal of public institutions and how are they used? Does the type of data determine the possibilities of its use or do those possible users determine how the data can be qualified? The answer to this question is not always clear.

Data with possible relevance for public institutions comes from many different sources and can be acquired in many different ways. Government organisations, citizens, businesses, researchers and other societal stakeholders all, knowingly and unknowingly, contribute to the generation of data, in open and closed formats, in big and small quantities, with and without structure, and of personal and non-personal nature.

The OECD's Working Party on Security and Privacy in the Digital Economy (SPDE) has been considering the breadth of attributes relevant to the discussion of the development of government-wide policies on the management and use of data but in considering the priorities for a DDPS this paper focus on three relevant attributes of data: big, open and personal.

Open data

'Open data' can be defined as "Data that can be freely used, re-used and distributed by anyone, only subject to (at the most) the requirement that users attribute the data and that they make their work available to be shared as well". We speak of Open Government Data when the data released according to these attributions are data produced by, or on behalf of, governments (Ubaldi, 2013[9]).

Big data

Datasets are becoming ever larger, more complex and available at higher speed. This development is generally referred to as 'big data'. The increasing digitisation of content, greater monitoring of human activities and the spread of the Internet of Things[1] are key drivers for the generation of big data (OECD, 2015[16]).

Personal data

Personal data means any information relating to an identified or identifiable natural person ('data subject') (OECD, 2013[17]). An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name,

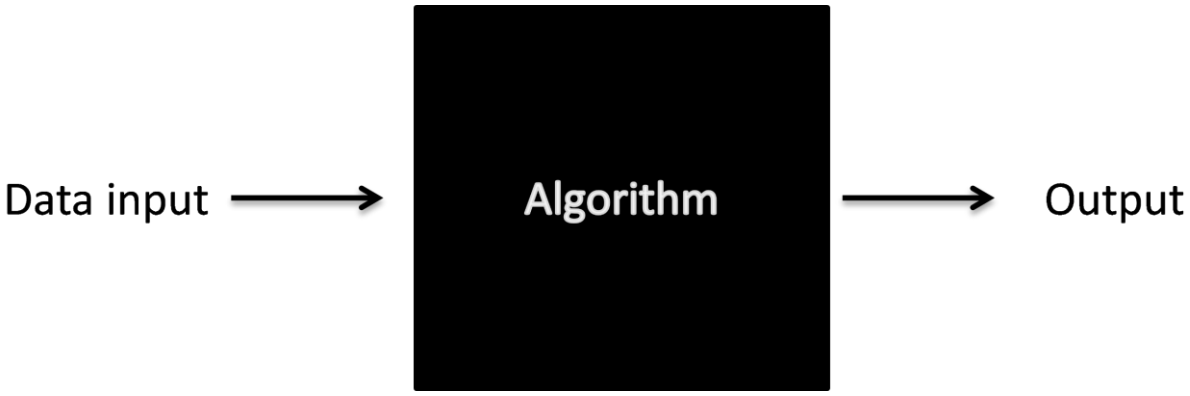
[1] Internet of Things refers to an ecosystem in which applications and services are driven by data collected from devices that sense and interface with the physical world (OECD, 2016[39]).

an identification number, location data or one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[2].

2.1.2. What use?

A strategic approach to the use of data necessarily entails a reflection on algorithms. "An algorithm is an explicit, precise, unambiguous, mechanically-executable sequence of elementary instructions" (Erickson, 2014[18]). While algorithms are as old as civilisation, their most commonly known manifestation is that of a computer programme (software). The use of computer algorithms in government is not new in itself, but the complexity of the algorithms and the range of application areas are vastly increasing, especially with the growth of machine learning, which are capable of continuously adapting themselves (Ubaldi et al., 2019[15]).

Figure 3. Algorithmic data processing



The increased volume, variety and velocity of data requires new techniques and tools for its processing and interpretation. These are generally referred to as 'data analytics' and can be used to infer relationships, establish dependencies, and perform predictions of outcomes and behaviours (Helbing, 2015[19]; Linturi et al., 2014[20]; OECD, 2015[16]). Four main categories of data analytics can be distinguished (Santiso and Roseth, 2017[21]):

- **Descriptive** analytics uses data to describe what has happened in analysing complex policy issues.
- **Diagnostic** analytics goes a step further by mining and triangulating data to explain why a specific policy problem happened, identify its root causes, and decipher underlying structural trends.
- **Predictive** analytics uses data and algorithms to predict what is most likely to occur, often by utilizing machine learning.
- **Prescriptive** analytics proposes what should be done to cause or prevent something from happening.

Increasingly the computers that perform data analytics are the machine equivalent of data scientists, possessing the capability to autonomously improve analytical methods based on what is learnt from their previous analysis. This development and mainstreaming of machine learning has been made possible by the revolution in access to sources of data,

[2] https://gdpr-info.eu/art-4-gdpr/

sometimes referred to as big data, that provides copious amounts of learning material for computers to use in training and improving their algorithms (Domingos, 2015[22]).

Governments are experimenting with how Artificial Intelligence can add value to policy making and service delivery, whilst data analytics are being successfully used to support taxation, census activities, and security and policing (Desouza and Jacob, 2014[23]; Ubaldi et al., 2019[15]). However, the overall promise of data in the public sector is largely unrealised. The public sector does not lack for data, but the "big data" hype and "open government data" movements have brought to light the insufficiencies in how public sectors handle the abundance of government data being collected or produced (Ubaldi, 2013[9]).

2.1.3. What actors?

There are six different actors involved in efforts to implement a data-driven public sector:

- the data subject[3]
- the data producer/provider
- the data controller[4]
- the data processor[5]
- the data analyst, and
- the public decision maker.

The **data subject** is the entity about which data is generated or collected. This can be people (individuals and groups), animals, objects (e.g. electricity meters, traffic lights, rooftops) or at times less tangible entities, such as organisations (public, private and not-for-profit) or natural phenomena (e.g. the sea, the air). Whether a datum can be linked to a specific person is an important factor in its reuse possibilities.

The **data producer/provider** is a source of non-personal data, not only for themselves as individuals or other individuals, but also about other data subjects. For example, citizens produce a wealth of data through their online interactions and mobile device habits that they consciously (or unknowingly) share with others and which can then be used for a purpose other than that originally intended by the data producer. For example, New York City police uses Twitter data as a risk indicator of criminal activity. The conditions under which citizen data can and should be available to public institutions, as well as the ethical considerations about the use of data for public decision-making are important questions to be addressed in a data governance strategy

Different actors in the digital government ecosystem can act as data producer on the same topic, thereby providing different perspectives on the same policy issue. For example as seen here in the case of crime and criminality:

- Public sector institutions, which produce official crime data
- Citizens, whose tweets might highlight and report criminal activity

[3] In this report, the term is used more broadly than for personal data.

[4] The data controller determines the purposes and means of the processing of personal data.

[5] The data provider processes personal data on behalf of the controller.

- Academia, where research investigates and collates data on the cause of criminal behaviour
- Businesses, in particular those involved in security, who record and report incidents
- Media, who may cover issues of crime informed by a variety of sources producing a new analysis for consumption by the public.

The **data controllers** are those who will ultimately bear the legal responsibility for ensuring that the data have been looked after. In government, this responsibility may remain with a public servant even if the activity of managing the data and keeping it secure has been contracted to an external hardware or software provider. Indeed, when exploring new forms of collaboration that result in outsourced data management, governments must ensure that procurement and commissioning activity not only reflects the appropriate level of responsibility for data management but also ensures that contractual arrangements allow for the movement of data from the incumbent to any new supplier as needed.

Data processor include external companies that may offer data management and analysis tools and critical computing resources, including, but not limited to, data storage servers, database management, analytical software and cloud computing resources.

Data analysts include those within government who perform data aggregation and analytical functions but also external companies, which provide these services to and for government. This also includes data visualisation services.

Finally, **public decision makers** who can be found at all levels of government ranging from ministers, policy designers, tax inspectors and street-level bureaucrats to budget and human resource managers. These different professions are empowered to make decisions that contribute to better policies, decisions and services. Their use of data should be encouraged in all aspects of the policy cycle, whether in initially understanding citizen and societal needs, designing responses to them, delivering policy and services to make a difference, or being accountable for its performance. In this way, data become an essential asset for governments to become forward-looking, inclusive, and responsive to citizens' needs whilst improving the efficiency and productivity of its processes, which can all contribute to strengthening public trust.

2.1.4. Whose data?

In the course of our daily lives and the use of our electronic devices, we create enormous amounts of data. When we shop online, use an app to track our workout, choose a route to drive to work, ask our home assistants for help (e.g. Google Home), or even simply move around with a phone in our pocket, we're creating data.

This proliferation of data beyond our knowledge, and often without our conscious consent, has prompted debates about ownership of the data we produce. For Jaron Lanier and others, it is clear that businesses and governments should recognise data as belonging to individuals (or users of devices or services). In this case we own our personal data and it should be treated the same as any other possession (Lanier, 2010[24]).

Martin Tisné presents an alternative view that "data ownership" is a flawed, counterproductive way of thinking about data, which does not fix existing problems and actually creates new ones (Tisné, 2018[25]). He claims that data in the aggregate can be fundamentally different from the individual bits that make it up. Data on their own are not very useful. However, when analysed in conjunction with similar data from thousands of other people to feed an algorithm or produce a composite, its utility increases exponentially.

He suggests that even well intentioned arguments about data ownership assume that if one regulates personal data well, one will get good societal outcomes but this cannot be guaranteed. This is why he argues that many of the problems about the unfair use of data cannot be solved by controlling who has access to it. As controlled access and re-use of data are seen as tools to minimise risks, he claims that even if a person denies consent to the use of "their" data, a public sector organisation can use data about other people to make statistical extrapolations that impact on the individual. For example, judges increasingly use algorithmically generated "risk scores" in making sentencing decisions and to predict the likelihood that a person will commit future crimes. If other people who are demographically similar tend to have criminal tendencies then such an algorithm may indicate a higher likelihood that the person will commit another crime and they may be treated unfairly and not on their own merits. Nevertheless, in such a situation no one can own their demographic profile or their own criminal record and refuse access to the legal system. Therefore, if an algorithm is judged unfair because it wrongly classifies an individual based on the biased or twisted data it was trained on, or because the person is an outlier, then recognising data ownership would not make data usage fairer.

Tisné, and several others, therefore suggest a different approach. Their argument is for a framework giving people rights to stipulate how their data is used without requiring them to take ownership of it themselves. The OECD's work on privacy and particularly the Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data are providing the basis for countries to develop their own approaches to safeguard their citizens' data along these lines (OECD, 2013[17]). The US Data Care Act, is a step in that direction. This new bill introduced in the US Senate in December 2018 proposes to hold large tech companies, specifically "online service providers", responsible for the protection of personal information in the same way banks, lawyers and hospitals are held responsible. The Act is designed to protect users' information online and penalise companies that do not properly safeguard such data. The bill would also protect personal information from being sold or disclosed unless the end user agrees, as it would establish three basic duties: the duty of care, the duty of loyalty and the duty of confidentiality (Ehlers, 2018[26]).

Finally, Tisné calls for the definition of a new paradigm for understanding what data is and what rights pertain to it if the intention is to forge an equitable 21st-century polity. He suggests that what is needed is a Bill of Data Rights that should include things like :

- The right of the people to be secure against unreasonable surveillance
- No person shall have his or her behaviour surreptitiously manipulated.
- No person shall be unfairly discriminated against on the basis of data.

Such a Bill of Data Rights would require a new model for data governance. In order to safeguard the rights it lays out there would need to be a new set of institutions and legal instruments that govern this new data-rights infrastructure and which protects and delimits those rights. The European Union's General Data Protection Regulation (GDPR) of 2018 represents a valuable example of what this could look like.

3. Opportunities

This chapter outlines the opportunities that can be derived from exploring the adoption of a DDPS culture within governments. It is broken into three sections.

The first section considers 'Anticipatory governance' or the way in which data can improve the capacity of government to act quickly when alerted to particular societal needs, mitigate the risks associated with emerging trends, forecast future needs and embed a more sophisticated approach to data in the policy development processes.

The second section looks at 'Design and delivery' and considers how governments can make greater use of data to understand problems and identify their users before better responding to their needs. This includes engaging the public with open approaches to communication and proactive efforts for generating data perhaps through crowdsourcing. It also reflects the shift from top-down assumptions about implementation to user needs led design approaches to services that can reflect contextual variations to target services and enable flexibility.

The final area of opportunities covered in this chapter are those associated with 'Performance monitoring and management'. The value of a DDPS approach is suggested as impacting on the productivity of the state in terms of efficiently planning and deploying financial, time, human and material resources. Finally, this chapter discusses how DDPS performance measurement can be harnessed to improve services and the integrity of government.

DDPS can increase public value and contribute to the well-being of nations in several ways. Governments are increasingly taking actions to enable themselves to leverage data as a strategic asset. For example the US President's Management Agenda (PMA)[6] released in March 2018, presents a long term vision to manage and use government data in order to modernise the Federal Government and improve the ability of agencies to deliver mission outcomes, provide better public services and steward tax payers money to produce better value (United States White House, 2018[27]).

The OECD framework for understanding the opportunities of a DDPS (Figure 4) identifies three areas in which data-driven initiatives are being developed to support the decision making process across policy areas and levels of government: *Anticipatory governance; Design and delivery*; and *Performance management*. This chapter will consider how governments can strengthen their capacity to use data for these purposes.

[6] For a full reading access: https://www.whitehouse.gov/wp-content/uploads/2018/03/Presidents-Management-Agenda.pdf

Figure 4. Opportunities of a data-driven public sector

Anticipatory governance	Design and delivery	Performance management
• Forecasting to proactively identify developments and future needs • Foresight to prepare for multiple plausible alternative outcomes	• Better predicting policy solutions • Engaging with citizens as co-value creators • Responding better to citizens' needs	• More efficient use of resources • Increase in resources • Higher quality and evaluation • Continuous improvement

3.1. Anticipatory governance

In a time of rapid transformation on an unprecedented global scale, governments need to significantly improve their capacity to anticipate, prepare for and navigate future change. Anticipatory governance refers to systematic efforts by governments to consider the future in order to inform policy decisions today. It allows governments to respond proactively rather than reactively, based on knowledge and evidence rather than experience and protocol. A data-driven public sector can strengthen its anticipatory capacities and support, complement and build on existing future-oriented approaches. Strengthening governments' capacity to anticipate citizens' needs and societal trends can help design responses that provide balanced opportunities for societal well-being and economic growth. There are two main forms of anticipation: forecasting and foresight.

Forecasting means using existing data and trends to try to predict the most likely developments and outcomes. Data-enabled predicting and modelling techniques may support governments to anticipate societal, economic or natural developments that are likely to occur in the future, but also allows for the development of early-warning systems, sentiment analysis of social media and real-time decision support systems (Höchtl, Parycek and Schöllhammer, 2016[28]) to better assess the need to intervene, design the appropriate policy measures and anticipate their expected impacts more precisely.

For example, aid plans regarding earthquakes could be activated upon early warning signs instead of after the fact, possibly preventing some types of damage, saving lives or limiting their longer-term damage. Additionally, data analytics can improve forecasting activities allowing governments to spot emerging needs and anticipate new ones based on observable trends in the physical world, as well as digitally. This in turn allows governments to recognise the opportunities available to them to shape policy and services to face the challenges of the future better. Furthermore, the content of those plans could adapt according to the anticipated sequence of events and impacts, as is being done for future disaster preparation in Japan[7].

Foresight, in contrast to forecasting, makes no attempt to predict the future but instead systematically explores multiple plausible versions of how the future could be different from expected, and then uses them to make policies more prepared and agile today. Data

[7] For more details about how the vast amount of information related to the 2011 Great East Japan Earthquake is being utilised for future disaster preparation by government, business and academia, see http://www.nhk.or.jp/datajournalism/about/index_en.html (accessed 1 September 2016).

analytics provide important inputs for foresight processes, which seek to identify emerging changes in the present which could grow into significant and disruptive developments in the future.

A data-driven approach can also help to identify key critical uncertainties where the outcome cannot be predicted based on today's knowledge. These critical uncertainties can be used to develop alternative future scenarios for the purposes of identifying future-fit policies that work in a range of different futures. Data-driven analysis can further help to spot when elements of such scenarios may be occurring, thus allowing for timely activation of contingency plans.

For example, in Portugal, funding has been directed towards projects that allow for the mitigation of risk. The country's *Roadmap to Innovation* has consequently identified work in the areas of health, employment and inspection. In particular, this trend spotting activity has been used to perform an analysis of the skills present in the unemployed labour force compared to the needs of the job market and subsequently identify those most at risk of becoming long-term unemployed and in need of targeted training.

The Mexican government founded a specialised data laboratory *Datalab* to stimulate the use and analysis of data for evidence-base policy development, implementation and evaluation[8]. One stream of activities is aimed at strengthening anticipatory governance and generating data-based predictions of, for example, populations at risk regarding diseases, zones with emerging environmental problems and future arising conflicts.

The role of data in terms of *Anticipatory governance* is also about the use of evidence to design policy interventions that respond to forecasted challenges. This is not about predicting whether a policy will be effective, but about how countries experiment with particular approaches or consider a wider spectrum of intervention in order to shape future policy interventions, informed by the availability of data. In this regard, it is interesting to reflect on the role which think tanks, the media and academia play in addition to the activities of government as there are a wide range of actors considering how to respond proactively to the needs of society based on data.

The data-driven approach is being embedded across many different policy areas. For example, in the case of monetary policy, the Bank of England has a dedicated Advanced Analytics unit, responsible for expanding the boundaries of data and data analysis at the Bank[9]. In Spain, predictive technologies are being used to support government decisions about research and development grants (Poel et al., 2015[29]). Some regions of New Zealand have used OVERSEER®, a national model for farm-scale nutrient budgeting and loss estimation to support environmental policy development (OECD, 2015[16]). The model calculates nutrient flows in a productive farming system and identifies risks of environmental impacts through nutrient loss, including run-off and leaching.

Data-driven anticipatory governance may also feed into the two other areas of opportunity for the data-driven public sector as presented in this paper: public policy and service delivery and performance management. Inclusive and user-driven approaches towards government service delivery can be strengthened through a forward-looking approach. The analysis of data on existing users' behaviour and characteristics can facilitate predictive

[8] For more details, see https://datos.gob.mx/blog/abrimos-convocatoria-datalab?category=noticias

[9] For more details, see http://www.bankofengland.co.uk/research/Pages/onebank/advanced.aspx (accessed 25 July 2016).

analytics on potential new users and/or emerging needs and problems when using existing services. Services adaptation may be offered proactively according to data on successful and unsuccessful use patterns. Additionally, data-driven anticipatory governance can contribute to performance improvement in allowing governments to speed up the decision-making process, revert or adapt decisions and better allocate and focalise public investments in priority areas.

3.2. Data-driven design and delivery of policies and services

Data can be used to better understand the dynamics of problems as well as identify the relevant stakeholders and their needs. This allows for new ways to engage societal stakeholders in policymaking and the development of services that respond to the needs of their users. Such an approach recognises that some of the data required for better insights must be generated by the public themselves in order to develop services which are driven by the needs and experiences of users, rather than the assumptions and beliefs of public servants.

3.2.1. Engaging society in policy and service delivery

Using data to foster civic engagement in public services and policies design can strengthen more people driven and participatory democracy (OECD, 2018[8]; Ubaldi, 2013[9]).

In the data ecosystem, governments interact with citizens and businesses carrying out the dual role of data provider and data consumer. Indeed, they contribute to the production evaluation, correction and mashing-up of data in ways that create value for others by guaranteeing the quality, usefulness and relevance of data. Moreover, data 'prosumption' redefine the relationship between the actors, creating a public value proposition where stakeholders and government can engage in more openly participative, proactive and empowering relationships in areas of policymaking, service design and community building.

Data analytics can support the process of debating policy options with stakeholders by providing detailed information about a policy problem in an accessible way. Methods such as interactive data visualisation help make conceptual issues real in ways that can be presented for different audiences and that provide specific entry points to engage stakeholders in contributing potential solutions. Additionally, by making data accessible and understandable it moves the conversation beyond the realm of experts and can increase levels of engagement by all, as highlighted in the United Kingdom's Open Policy Making Manual[10].

Besides facilitating traditional methods of stakeholder engagement, data analytics can help governments analyse the wealth of input coming from crowdsourcing initiatives. Automated analysis of data regarding citizens' opinions and/or behaviours through such channels as public blogs, social media or public consultations provides insight on the public reactions towards certain policy problems.

In Paris, the municipal division for the environment crowdsourced data on potential locations for new green spaces before launching a revegetation campaign, which allowed

[10] For more information, see https://www.gov.uk/guidance/open-policy-making-manual/getting-started-with-open-policy-making (accessed 1 September 2016).

citizens to apply for a revegetation permit and manage a new green area on their own (De Feraudy, 2017[30]). In France itself, the government is supporting a project using crowdsourcing and data analytics to understand the potential for French rooftops in supporting environment-related initiatives such as the installation of solar panels or rooftop vegetation[11]. Citizens classified rooftops in an open geographic database, with that data providing the training material for a machine-learning algorithm to analyse 48 million referenced rooftops elsewhere in the country. The contribution of citizens coupled to data-driven analytical techniques generating data of sufficient quality and at a large enough scale to be useful for public policymaking.

3.2.2. Data to improve public service delivery

Data analytics enable a closer working relationship between policy design and service delivery activities with a resulting shift from top-down implementation of public services to a user need led approach to design and delivery, based on an end to end understanding of a particular service journey, which can consequently increase its reach and effectiveness.

Policy problems increasingly need the simultaneous attention of specialists from different domains[12]. For example, ensuring the well-being of ageing populations involves collaboration between specialists on health, housing and social services, amongst others with the recognition that actions in one domain affect others. Combining data produced or collected by different specialised public organisations allows for the development of integrated, interagency policies designed in response to the end-to-end need of a citizen, not just the part which one organisation handles.

Furthermore, this allows for a more granular analysis of policy problems, producing the insights that lead to contextual variations in policy design and implementation, which can boost policy effectiveness, lessen unnecessary burden on citizens and allow government to better target scarce resources. For example, policing could be focused more precisely on those areas with high criminal activity (Höchtl, Parycek and Schöllhammer, 2016[28]).

The increasing ubiquity of mobile devices and social media platforms means public authorities have a wider range of routes to gather valuable information from citizens related to their daily lives as well as their needs, preferences and behaviours. Citizens are actively reporting issues such as road quality and earthquake situations as well as suspicious behaviour, allowing governments to take action in response to the information shared by the public. Moreover, by communicating how that data have been used in a way that is accessible to the public, public authorities can have a positive effect on citizens' sense of achievement and their level of satisfaction with government services (Morabito, 2015[31]).

This is important because the trust relationship between citizen and government is key to the success of DDPS approaches. The aggregation of citizen data from different public institutions allows governments to build up a more complete view of the needs their citizens have. Such a view may be viewed negatively by citizens as indicative of a 'surveillance state', even if the intentions are noble. A DDPS is better able to respond to the actual needs

[11] For more information on the Opensolarmap Project, see https://www.etalab.gouv.fr/opensolarmap

[12] Similarly, in academia, more and more research problems are recognised as requiring an interdisciplinary approach with collaboration between researchers from different disciplines. For example, protecting privacy by embedding it into the design specifications of technologies, business practices, and physical infrastructure (Privacy by Design), requires the involvement of legal, technical and business specialists.

and personal situations of their citizens by enhancing the use of data in the design and delivery of services to produce results driven by an understanding of the user and their needs, based on data willingly provided by the citizen.

Being able to automatically include the data citizens produce or offer into policymaking and service design cycles simplifies the challenge for public institutions of knowing how their citizens use services, and ensures that designs can respond accordingly. As a result, government services can continue to improve, responding to the rhythm of people's lives and continuously evolving to meet their needs and expectations.

An example of such a service can be found in Mexico, where the Ministry of Social Development is building a new system in which household, beneficiary and geographic data are combined to target social service programs to those who need them most[13]. The new system may allow for identification of individuals who qualify for, but have not used, particular programs, prediction of household needs potentially informing the design of new social service programs, and detection of people who have under-reported income to receive assistance.

Data exchange can also facilitate new forms of international cooperation in addressing cross-border issues like illicit trade and improve service delivery in transport and migration services thanks to the automated capturing and processing of data related to citizens and businesses' needs and behaviours[14]. For instance, the automated exchange of basic business register data are expected to facilitate and stimulate business development between Estonia and Finland (OECD, 2015[32]) whilst the European Union's eIDAS standards are designed to allow citizens of one member state reuse their digital identity when accessing the services of another.

3.3. Data-driven performance monitoring and management

A DDPS recognises that ongoing performance monitoring and management can be enhanced through the application of relevant data. In addition to improving service design and delivery, this may also benefit from learning about how data are being used within organisations in order to make services and products better. This then feeds back into the policy design and service delivery activities of a country to provide holistic, iterative, approaches to the role of data in responding to the needs of a country. To do this requires an effort to make sure that those involved in performance monitoring and management are aware of relevant data, and have access to it.

3.3.1. Improving internal public sector productivity and efficiency in the use of resources

Governments can achieve considerable benefits and productivity gains through better allocation and management of money, time, human resources and materials from sharing data and embedded data analytics in their processes. This may allow government activities to be streamlined with a resulting reduction of operational costs. Furthermore, strategic, tactical and operational decision-making processes can run more smoothly because, if the

[13] For more information, see https://dssg.uchicago.edu/project/enhancing-the-distribution-of-social-services-in-mexico/

[14] For example 'Project Crocodile' regarding the governance of cross-border transport data, http://2016.itf-oecd.org/awards

required data is already available, less time is needed to request or calculate the necessary information.

Indeed, data are a critical resource for all those in government involved in planning, and in particular those who are looking to improve the internal productivity and overall business performance of public organisations. DDPS has the potential to boost public sector productivity not only at the micro (organisational) level, but also at the meso (sector) and macro (whole public sector) levels too. As an example, in Chicago, publicly available food safety data is used to prioritise restaurants for inspection, thereby improving the efficiency of inspections and ensuring that citizens are protected from unhygienic establishments (Martin and Begany, 2016[33]).

Financial resources

Data-driven governance offers significant room for improving the efficiency of how public administrations use their financial resources (OECD, 2019[34]). The application of data offers new ways of addressing the tax gap, lowering procurement costs and detecting erroneous, illegal and unethical behaviour in ways that reduce fraud, corruption and error, and increase government revenues and reduce expenditures. For example, in the United Kingdom, the Digital Economy Bill makes provision to enable data sharing for the detection and prevention of fraudulent activity in order to reduce the financial loss to government that is happening each year[15].

Risk managers and auditors, among other functions in government, can benefit from data-driven approaches that facilitate risk assessments, audits and audit programming. For example, risk managers or procurement officials can match databases to identify potentially high-risk contracts, such as where a bidder has previously received sole-source contracts within specified advertisement periods of the tender. These data-driven approaches do not remove the need for the skilled professionals but complement them and provide higher confidence, more quickly. The Brazilian auditory body, the *Tribunal de Contas da União*, has identified similar benefits by using data analytics, including statistical methods, image analysis and geo-referencing techniques, to reduce the burden for auditors assessing the cost of public works.

Time resources

Operational efficiency gains can be expected from reducing the time it takes to process applications in the back and front offices of government as well as adopting design approaches that reuse existing information following the once only principle to reduce duplication and errors. For example in Estonia, tax returns are provided to citizens two days after the submission of tax forms, which are fully prefilled based on data coming from public and private sources (Tamkivi, 2014[35]). Consequently, whilst citizens do save time the greater public value accrues to an institution whose use of this data-driven approach leads to greater operational efficiency.

Human resources

People analytics open up new doors to recruit, retain and manage the performance of civil servants, as well as understand how to improve their working environment to enhance

[15] For more information see https://www.gov.uk/government/news/improving-the-way-government-shares-data-to-transform-public-services

productivity (Isson and Harriott, 2016[36]). People analytics can also help to understand and benchmark the performance of strategic HR systems and management in order to maximise the impact of human capital. Such an approach aligns with the efforts already made in many OECD countries to pursue evidence-based management of employee. In Mexico, the Ministry of Energy is using a predictive workforce planning and analytics model to identify current and future talent and skills gaps in critical (Gerson, 2015[37])oil and gas occupations over a 10-year horizon (Deloitte, 2016[38]).

Material resources

As government assets, such as traffic lights, buildings and vehicles begin to be transformed into smart assets connected to the Internet of Things, automatic and predictive monitoring becomes possible. Data about the real-time status of such assets can be used to better plan maintenance with the least disturbance to service continuity as possible (OECD, 2016[39]).

Data analytics applied to the use of resources in the delivery of services and the completion of tasks can highlight inefficacies that enable the better management of available resources of all these types. Moreover, an understanding of demand for services themselves will provide a clearer view of the service delivery landscape in a society, allowing for better targeting of specific channels to specific segments of the population to ensure no citizens are left behind. A proactive public service that can intervene more quickly to cope with unforeseen demand by reallocating public servants or adapting opening times may generate considerable cost savings and increase the satisfaction of the public. For example, the open data application *Streetwiz* helps Los Angeles city departments to plan construction work in order to avoid having too many ongoing projects in one area that may impede traffic and liveability[16].

Nevertheless, concrete numbers and estimates about the productivity and efficiency gains that result from the better use of data by the public sector are scarce. Certain predictions have been made with McKinsey Global Institute calculating potential efficiency gains at EUR 250 billion per year within the European Public Sector (McKinsey Global Institute, 2011[40]) and between USD 35 and 95 billion per year in the United States by 2020 (McKinsey Global Institute, 2013[41]). Another indication of the potential for productivity gains can be found from the experience of the private sector, where companies defined as being 'data-driven' were on average five per cent more productive than their competitors (McAfee and Brynjolfsson, 2012[42])[17].

Box 1. The role of data analytics in the fight against fraud

France

Several public institutions in France are taking a data-driven approach to assessing fraud risks and adapting their inspection practices accordingly.

For instance, in 2011 the national public institution responsible for social benefits, *la CNAF*, identified that the wealth of data they held could be used for more than

[16] For more information, see http://datasmart.ash.harvard.edu/news/article/location-is-driving-the-value-of-l.a.s-open-data (accessed 30 August 2016)

[17] The authors state that this performance difference remained robust after accounting for the contributions of labour, capital, purchased services, and traditional IT investment.

administrative purposes. The analytical model they developed the risk assessment of fraud was then used to better target inspections. The success of the new approach was not only the quality of the model and its results but the active participation of inspection officers at the local level who were persuaded that a data-driven approach wouldn't replace their expertise on the ground but would provide support to them in achieving higher catch rates, and more time to focus on unusual cases.

Data mining to inform the risk analysis that decides where inspections should take place has increased the rate at which fraud is discovered. This results in increased tax revenues and lower expenditures on unemployment and social benefits as well as saving time, which inspectors can then dedicate to investigating more unusual cases that cannot be resolved by machine, thereby increasing the success rate in this category too.

Source: Interviews with various stakeholders in the French government

Australia

The Auditor-General of Western Australia used a variety of data-analytic techniques to audit payroll and other expenditure data for potential fraud and corruption. The audit team downloaded and analysed 4 million transactions from 12 state agencies, totalling over $7.5 billion (AUS), which occurred during the period 1 July 2014 to 30 April 2015. The audit team tested data to identify a range of possible risk indicators, including:

- supplier invoices that may have been paid twice;
- officers approving payments that were above their delegation limits;
- 'splitting' of invoices into smaller amounts to circumvent system delegation limits;
- purchases from suppliers associated with agency staff, which may indicate undue;
- favouring of a supplier;
- government purchasing cards used while the relevant officer was on leave;
- unusual payments or payments with inadequate descriptions, which may indicate fraud;
- payments approved by only 1 officer;
- payments to non-approved suppliers;
- large payroll payments which may indicate overpayments;
- payment of allowances beyond their approved period; and
- employees paid after they ceased work at the agency.

Data analytics allowed for the application of audit procedures to 100% of transactions in some instances, as well as sampling of identified anomalies in others. The audit led to the identification of errors, including overpayments, and the need for improved controls at half of the agencies tested.

This broad assessment and systemic improvement to the control environment would not have been possible without the use of data analytics. The auditor made recommendations to agencies to periodically review and update policies and procedures, periodically review the adequacy of controls, and monitor compliance with policies, including thorough analysis of transactions and databases. The Auditor-General also recommended that

agencies consider the use of data analytics to identify fraud or errors after they implement a new system or process in order to identify gaps in new or revised controls.

Source: Western Australia Office of the Auditor General (2016), "Audit of Payroll and Other Expenditure Using Data Analytic Procedures," https://audit.wa.gov.au/wp-content/uploads/2016/05/report2016_06-DataAnalytics.pdf

3.3.2. Strengthening policy monitoring, evaluation and continuous improvement

A DDPS fosters an environment in which data about real time policy interventions is available. Policymakers no longer have to wait for monthly or quarterly updates in areas such as migration, unemployment and demographics because the data they need is becoming available, and accessible, with a higher frequency. Consequently, this provides better insights into the policy process and enables quick policy adjustments in the short-term with a resulting increase in accountability and continuous improvement in the mid- and long-term.

Those responsible for policy can more easily establish whether their policies are having the desired effect and, if this data is made available as open government data, so can other stakeholders. As a result, policy evaluation can turn into an open, ongoing process rather than an internal, snapshot moment. Data analytics can be especially powerful in identifying the adverse effects of a decision or a policy and being able to respond accordingly (Höchtl, Parycek and Schöllhammer, 2016[28]). Whilst the monitoring of performance might be prompted by a top-down desire for oversight and reporting on delivery, a DDPS is interested in how those insights can be analysed and, crucially, applied in improving performance based on a deeper understanding of the needs of the organisation and its users.

In the United Kingdom, one of the early activities of the nascent Government Digital Service was collating and publishing a list of all transactions with their associated costs and volumes. This made it possible to understand some priorities for transformation and informed the first phase of exemplar digital services. As their approach matured this turned into a platform facilitating the collection, publication and visualisation of 4 Key Performance Indicators for each of those services in real-time (https://www.gov.uk/performance). Alongside the technology, the Government Service Standard mandated teams to publish and understand their performance data. This meant there was not only a publicly visible resource for stakeholders highlighting performance but, crucially, a unifying pan-government culture of service level performance.

In Argentina, the open government data portal is used to showcase examples of data-driven journalism that have used, and then expanded upon the published data. In one case third parties writing for *La Nacion* have consumed openly available data on public servants' declarations of assets which then becomes an additional input for those within the Ministry of Justice's Anti-corruption Office who monitor the declarations of interest from public servants (OECD, 2017[43]). This initiative illustrates the potential for open government data to be analysed by external sources and generate further value not only for non-governmental users and the general public but for the internal activities of the public sector itself.

Supreme Audit Institutions (SAI), who traditionally exercise external oversight of government accounts, have recognised this potential to innovate and expand their activities. In 2016 the International Organisation of Supreme Audit Institutions, the umbrella

organisation for the SAI community, established a Working Group on Big Data to address the challenges and maximise the opportunities in this area with the result that SAIs are accessing new data sources to enhance their capacities and audit work. In the Netherlands, the Court of Audit uses data from the Dutch secret service, Tax Office, and Ministry of Defence to advance a variety of audit objectives (Janssen, 2015[44]).

SAIs make use of the data they can obtain and analyse to detect potential and occurred fraud and corruption, conduct performance audits and evaluate outcomes, and generally offer insight and foresight about developing trends or systemic issues facing governments (Government Accountability Office, 2015[45]). For instance, the Auditor-General of South Africa conducts reactive and proactive analysis of fraud risk indicators in financial statements to understand the extent and impact of a perpetrated fraud. Similarly, several SAIs, including the U.S. Government Accountability Office, the Netherlands Court of Audit and the Chinese National Audit Office, are developing data analytics capacities to improve insight and foresight auditing in different sectors. In China, data analytics and big data are being used to track the lending and asset quality of commercial banks, and spot financial and operational risks. The results of such analysis provides direct input to improve the integrity, effectiveness and inclusiveness of public policymaking and service provision (see Box 2), thereby contributing directly to organisational learning and performance improvement. As such, SAIs are important players in building data-driven public sectors.

Box 2. Innovative performance monitoring by the SAI in India

In addition to preventing and detecting fraud or corruption, data analytics is being used in other areas of SAI work. For instance, the Controller and Auditor General of India is applying data analytics to audits of the social security programme, to improve its capacity to protect citizens, particularly vulnerable populations, from unforeseen circumstances and financial risks.

The conventional performance audit of the social assistance programs of Old Age Pensions, Widow Pensions and Disabled Pension schemes was designed to identify improper payments to ineligible persons. However, by introducing an innovative use of data, the SAI auditors were able to verify whether eligible beneficiaries had been excluded, producing a performance audit that was more thorough and insightful.

In addition, in February 2016, India's SAI established a new Big Data Management Policy, which included the categorisation of data sources into internal (that is created and maintained by the SAI) and external (that is available from audited entities or in the public domain). This policy includes a strict protocol to ensure data is authentic, relevant and usable and sets out guidelines for data access arrangements with external sources. Finally, the Indian SAI established a monitoring group to oversee the implementation of this new policy to optimise the potential for new data analytical technology.

Source: Big Data Analytics in SAI India (SAI India, 2016[46])

4. Challenges

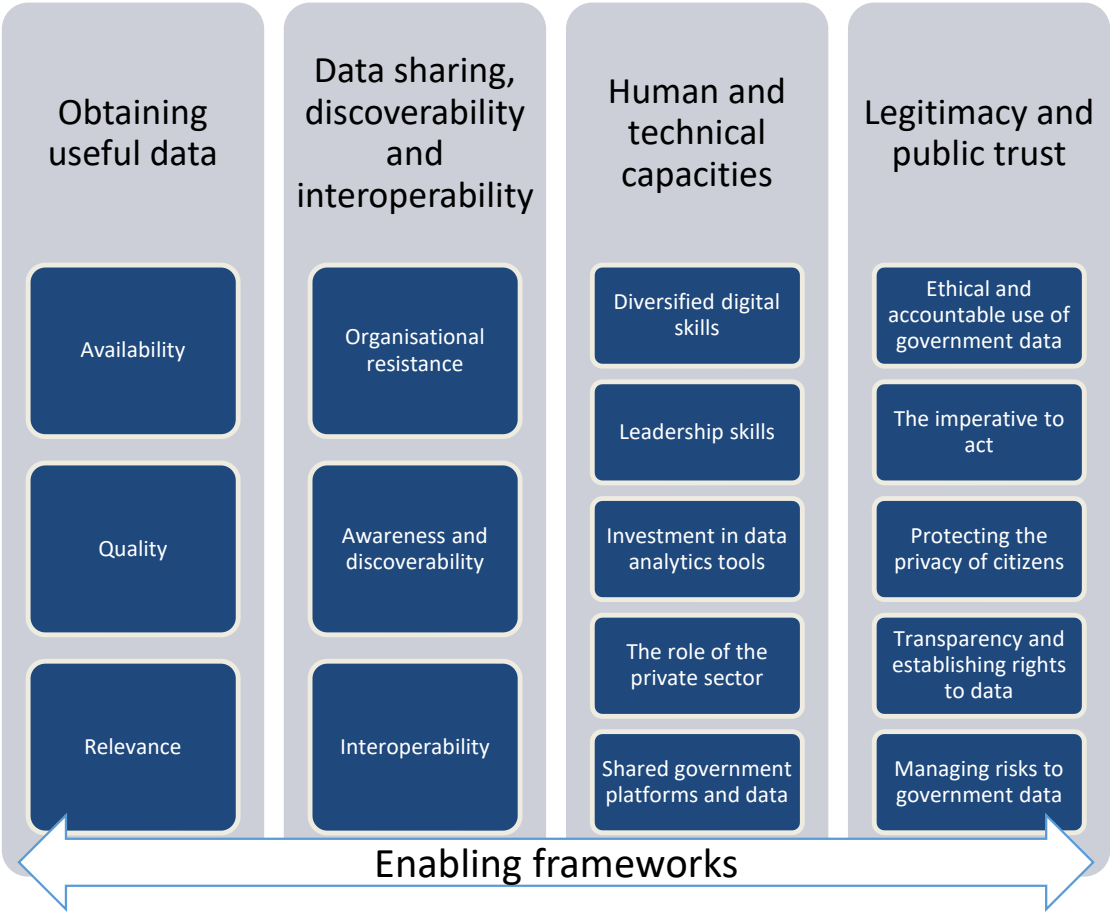
Having explored the context for a DDPS, defined some of the important concepts and considered the opportunities it provides to governments, this chapter will look at the Challenges facing governments wishing to implement a DDPS approach.

The first area is that of 'Obtaining useful data' and raises questions of the availability, quality and relevance of the data which governments access. The second area relates to 'Data sharing, discoverability and interoperability' and involves the consideration of organisational resistance, awareness of the data that is available and the underlying challenges of interoperability between sources of data. The third area relates to 'Human and technical capacities' and looks at the needs for digital skills, digital leadership, the tools themselves, the role of suppliers, and how governments can better collaborate on data related issues. The fourth area concerns 'Legitimacy and public trust' and explores the ethical, privacy, transparency and risk challenges that governments may face.

To build on existing data-driven initiatives of governments to improve policymaking, service delivery and organisational management and successfully realise the potential of a DDPS, it is essential for governments to create the necessary institutional conditions to respond to the challenges they face. The more governments advance in the definition and implementation of the data-driven vision and agenda the more informed their understanding of the existing challenges. The most frequently mentioned challenges are the barriers to data access, sharing and use which are caused by limited or difficult collaboration within the public sector. Additionally, privacy and security considerations can lead to the adoption of protection measures considered legitimate protection measures. This can be the result of patchy and complex interagency agreements, a lack of common approaches or consensus on legal frameworks or procedures related to data identification, management and use, or inadequate data interoperability or data standards.

The success of governments to respond to the challenges identified in Figure 5 will determine the extent to which they benefit from the opportunities of a data-driven public sector described in Chapter 3.

Figure 5. Challenges in fostering a data-driven public sector



4.1. Obtaining useful data

The initial challenge for successful implementation of DDPS is in having access to data in the first place. A DDPS can identify, collect and obtain useful data for policymaking and service delivery, especially in support of anticipatory governance activities so that the quality of predictive modelling is reliable.

4.1.1. Data availability

The effective modelling of societal trends and training of machine-learning algorithms requires representative samples of data that also reflect and incorporate the experience of marginalised and hard-to-reach groups, even though this can be an issue with data availability. Challenges with low or irregular data availability can present a problem to ensure that DDPS driven policy interventions reflect the needs of those they will affect most. For example, those living in rural areas or with low incomes may be less well-connected to conventional online channels and consequently under-represented in modelling that relies on data alone (Mickoleit, 2014[47])

Therefore, an inclusive DDPS must pay attention to the appropriate methods of data gathering for groups that might otherwise be overlooked and consider the biases that may

be introduced into the sample of data being used. It is important for those involved with designing and evaluating policy interventions to continually ensure that they are using the most relevant data sources and to be exploring how they might increase the quality and feasibility of any data gathering processes.

4.1.2. Data quality

Every day governments collect and generate huge amounts of data. However, whilst it is not in short supply quality can be questionable. It can contain fraudulent untruths, accidental errors in transcription or misunderstanding and may be held in a format that is difficult to process. The quality of data very much depends on the purpose of its use. Thus, whilst governments are increasingly making data available, those who would analyse it tend to spend the majority of their time on data cleansing activities according to their needs (Bulger, Taylor and Schroeder, 2014[48]). Furthermore, innovative efforts to improve the lives of citizens can be completely blocked by the quality of data as seen in the experience of the Smart Citizen Kit in Amsterdam, which aimed at collecting environment data with affordable electronics and sensors and the help of a large group of citizen (Waag Society, 2014[49]). Results show that the quality of data collected was not good enough due to some major technology issues; therefore, the social impact of the experiment could not be conducted. This presents a significant challenge for the implementation of the DDPS as well-structured data is essential for effectively responding to a policy need in a data-driven fashion.

Therefore, if data is the fuel of the DDPS, it is imperative that its quality is guaranteed. This means governments must consider ways of supporting its data collection efforts and policymaking activities across a variety of dimensions (OECD, 2018[8]; Ubaldi, 2013[9]). These include:

- accuracy and precision
- comprehensiveness and clarity
- consistency and integrity
- design of metadata
- completeness
- uniqueness and relevance
- timeliness
- validity

Providing an effective response in these areas creates a responsibility for governments to define ways of connecting with different user communities to understand what data quality means to them, manners of measuring and reporting on the levels of data quality. Some governments may choose to create one standard while others may consider it more appropriate to adopt different approaches depending on the different purposes of data use.

If the quality of data does not meet the required standard by the specific user group, because it is either incomplete, not made available in a timely manner, out of context or lacking in integrity, the decisions they lead to can potentially have an undesired impact or negative consequences for the citizens it is being used to help (OECD, 2015[16])). Upfront investment in the design of data collection and data gathering methods to secure the availability of quality data can help to foster collaboration as a platform, and ensure that less effort is

required by citizens and government either initially, or later on (OECD, 2018[8]). As an example of how flawed data collection processes can lead to later issues, citizens in the Netherlands have found that it can take up to five interactions with public institutions to correct mistakes[18].

Nevertheless, it is possible to explore ways of offsetting issues of low quality data. The use of AI and machine learning specifically can help identify data that deviates from established levels of quality or expectation but in order to be successful the initial training data needs to be reliable. One of the ways to increase the quality of data, and provide the basis for algorithm-based decision making, is to open things up to wider scrutiny and use a crowdsourcing approach that encourages multiple eyes to examine parts of a dataset to confirm the quality and correct any errors.

However, these approaches must also be critically examined to ensure that they themselves do not unintentionally undermine the quality of data by introducing new issues. It is therefore crucial to have adequate governance of data correcting activities.

4.1.3. Data relevance

A further issue in establishing a culture that can support DDPS is the relevance and utility of data that is being accessed once it has been collected. This requires that governments consider how they design data collection methods, and how they conceive of the questions they are seeking to answer.

It is impossible for governments to measure everything. However, it is essential that those who wish to understand the performance of a situation and respond to those insights in order to make things better, identify effective means of measuring a given situation.

This is particularly important when data is being exchanged and then reused. In these circumstances data must be correct and of reliable quality but it must also be relevant to the case in hand. For example, when collecting sensor-based data on stroke patients' movements, it would be most appropriate to consider at the design stage how the collected data will be useful to caregivers and allow for transformation of care to take place (Klaassen et al., 2017[50]).

Sometimes it will not be possible to participate in the design of the original data collection meaning that those who consume the data will play a role in interpreting and filtering its insights. One of the increasingly prevalent ways in which this may influence the outcome of a DDPS intervention is in the application of Artificial Intelligence (AI) and the training data selected as the basis for any machine learning. Selecting the appropriate data for a particular application and recognising that it may well contain biases that would affect the final outcome is an essential first step. For example, there is a risk that the concerns of a vocal minority might be overemphasised if governments analyse social media data to inform the policy agenda (Lazer et al., 2014[51]).

4.2. Data discoverability, sharing and interoperability

Having established criteria by which data is made available, a DDPS must consider how to approach the ongoing use and application of that data within its public institutions by considering how data can be found (discoverability), made available for others to use

[18] http://www.rtlnieuws.nl/geld-en-werk/fout-in-gegevens-overheid-1-op-5-moet-minstens-5-keer-bellen

(sharing), or processed and reused (interoperability). Letting data flow across organisational boundaries faces a number of cultural, technical and procedural challenges.

4.2.1. Organisational resistance

It is a recurrent theme in the digital transformation efforts of governments to secure visionary and effective leadership for a future way of working that is not overly constrained by experiences of the past. For many, the approaches of a DDPS in the collection, sharing and reuse of data are an uncomfortable deviation from existing processes and methods. Some people will cite the loss of ownership as an objection to sharing whilst others will fear that data will be misinterpreted, compromise confidentiality, or increase their own workload (Martin, Borah and Palmatier, 2017[52]).

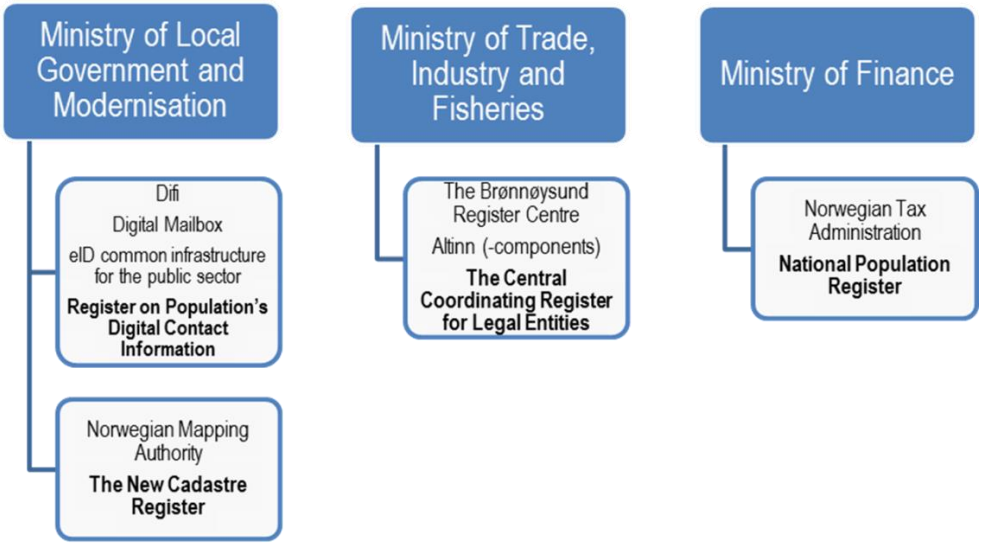
Such organisational resistance and reticence is a cultural challenge that must be overcome. As with other challenges facing those who would make progress in the implementation of digital government, the starting place for this issue relates to siloed thinking and a protectionism of delivery activity and data that creates the conditions that lead to conflicting interests between government organisations and individual departments.

It is therefore crucial to establish an understanding across departmental boundaries and organisational divides that there is an 'all government need of data'. Failing to achieve this consensus will pose a significant threat to data sharing efforts, and consequently produce duplicated activity for citizens and limitations on the extent to which governments can embrace DDPS practices.

Within the European Union, the challenge of data sharing within government has been the focus of efforts to establish the 'once-only principle' since it was endorsed by EU ministers in the Ministerial Declaration on eGovernment in 2009 (European Union, 2009[53]). These efforts aim to establish a baseline for interacting with government where those dealing with government should expect to only provide information once, with it being simply reused in subsequent circumstances. Many of the countries in the EU have embraced the principle but, as implementation was the responsibility of each country, their responses have differed.

This legislation based approach may prove the most effective in establishing sustainable cross-government practices of data sharing. In Norway, the country has developed a mature network of basic data registries, with clear definitions about who is responsible for each of the key tasks associated to the government data management, ownership and value chain (OECD, 2017[10]). The existence of clear responsibilities in relation to the management of the registries, of which the Brønnøysund Register (entity that manages several public sector registers) is a good example, provides the perfect backbone and foundation for a solid governance to develop a DDPS (Figure 6).

Figure 6. Norway's base registries



Source: Digital Government Review of Norway (OECD, 2017[10])

However, organisational resistance is not just an issue of cultural or environmental factors, there can be important, legitimate, reasons for caution about the use and reuse of data from a legal point of view in terms of data control and financial impact. In cases where governments have a financial overhead for providing the technical and organisational infrastructure to gather and then process data, that data is understandably seen as an asset from which income can be generated. Changing the culture to make that data available without a fee structure requires a strong business case that argues and proves the wider value such a move would offer and may need governments to revisit the funding arrangement of the original institution.

4.2.2. Awareness and discoverability

Creating the conditions under which those in government are more open to sharing data across organisational boundaries is an important enabling step in supporting a DDPS but governments also face challenges in the discoverability and cataloguing of the data throughout the public sector. There is not only a need for agencies to be aware of the data they are responsible for collecting but they must also consider the data-altering processes they are managing when it comes to the lifecycle of their data assets. This is because the rest of government must have a reliable source of information about what data is available, and the conditions under which that data might change over time.

By establishing this pattern of data management within individual parts of government it becomes possible to aggregate disparate sources into a catalogue, or index, of available data. Ensuring that this is readily discoverable for service teams, policy designers and government suppliers will maximise the opportunities for realising the tremendous potential of the data. As this simplifies the mapping of data access and flows throughout government it makes it possible to identify, and address, historic duplication of data collection and storage. This is desirable as not only are these issues a drain on government

efficiency they are a burden on citizens when asked to repeatedly provide the same information.

4.2.3. Interoperability

Alongside the culture within government organisations and the ability for people delivering services to know about and access available data effort must also be placed on the usability of data across the public sector. Therefore, an important question for all those wishing to implement DDPS is how to guarantee the interoperability of data.

Some of the considerations that this produces are **technical** by nature, relating to the system that produces the data, the formats in which it is published and questions around its quality; and others relate to **semantics** in terms of metadata or, where a country has multiple official languages, challenges of multiple languages. However, the section below underscores important considerations in this respect that relate to the **governance and leadership** needed for the successful creation of consensus around an approach to interoperability that enjoys the support and adoption of the public sector.

Interoperability is a critical issue at the national, and regional, level to ensure that the public sector can tackle issues as a whole and maximise the potential offered by DDPS. However, the sharing of data across international borders is increasingly important as countries explore cross-border services. As such, it is important for countries to participate in efforts to establish and adopt universal data standards if possible. Within the European Union, 47 concrete recommendations have been developed as part of the European Interoperability Framework (EIF) to support member states in protecting efforts at establishing national, and international interoperability. Furthermore, in Korea, the government is adopting an approach that enhances Enterprise Architecture for managing data resources both domestically and internationally.

4.3. Human and technical capacities

This paper has so far provided several arguments underlining the relevance of data as a strategic asset for the public sector to improve policy design, service delivery and the overall and operating of the machinery of government. Yet data will not produce the value without the help of human and technical capabilities. Therefore, to take full advantage of DDPS it is necessary to identify how those capabilities can be most effectively mobilised.

This section considers five areas where governments must respond to challenges around these capabilities in order to maximise the value of DDPS. They are the breadth of digital skills required across the public sector; the role of leadership in particular; the necessity to invest in data analytical tools; the role of the private sector in providing both people and technology; and how shared government platforms can be used. Lifelong learning is also a debatable attitude, but will not be discussed in this report.

4.3.1. Diversified digital skills

Governments are already responding to the skills challenge that results from the impact of digital transformation on the way in which public services are designed and delivered. In order to be successful in embracing DDPS, skills associated with data literacy, communication and management must also be considered (Grundke, R. et al., 2018[54]). In responding to these diverse skill needs there are several questions that governments must answer:

- What new skills are required from data scientists in government organisations, from policymakers working with them, and from back-office and frontline civil servants who directly work with data for service delivery?
- How can frontline professionals be empowered to act on that data to rethink policy and service delivery to respond and adapt better to real time information?
- How can we get them to be more data-based in their decision-making approaches in the first place?
- How can public officials learn to extract meaning from data by correctly identifying the policy problem to address and considering relevant factors to design potential solutions?
- How can they transform this into information and knowledge to inform decision-making?

It would be a mistake to consider that the effective use of data is only the responsibility of those with 'data' in their job title, whether as data analysts or data scientists. Fully embracing a DDPS approach will have a broad impact that requires system-wide thinking and cross-government support for decisions about investment and policy direction. As such, countries must ensure that an understanding of data's potential and an awareness of the issues in its management are a core skill for public servants (Grundke, R. et al., 2018[54]). Several countries have developed dedicated skills strategies to accompany public servants in the digital transformation, with the United Kingdom considering a range of skills, including a particular focus on data (see Box 3).

Box 3. Digital skills in the United Kingdom government

The United Kingdom's Government Digital Service (GDS) Academy provides training for public servants to equip them with the digital skills necessary for transforming public services. Courses are delivered from five locations across the country and focus on teaching public sector professionals how to embrace the changes brought about by digital transformation and to work more efficiently and creatively.

Courses are available to civil servants, local government employees, devolved administrations and other public servants across all grades and levels of seniority. They cover such general topics as 'digital and agile awareness' and 'research and design in government' as well as profession specific training to product managers, delivery managers and analysts.

Many of these courses incorporate a module that discusses the importance and potential of applying data to the design and delivery of public services. This ensures that those who are being equipped to explore the opportunities of digital transformation in the UK government are aware of the challenges and opportunities to consider in terms of the data they are responsible for collecting or processing.

The efforts of the Academy is complemented by a specific Data Literacy programme aimed at senior leaders and managers as well as civil service wide training opportunities via its internal Civil Service Learning platform.

Furthermore, the Data Science Accelerator is a capability building programme targeted at the needs of those in specifically analytical roles. It is a joint project between GDS, the

Office for National Statistics, Government Office for Science and Civil Service analytical professions (statistics, economics, operational research and social research).

4.3.2. Leadership skills

It is important for governments to explore interventions that will help public servants at all levels build confidence in the possibilities of DDPS. However, it is critical that these efforts are complemented by government leaders who demonstrate an appetite to adopt a new way of working and are willing to use their influence to remove obstacles and provide support. To understand how best to equip these individuals governments will need to resolve several questions including:

- How can leadership be developed to overcome resistance and foster support?
- How does leadership change in a data-driven public sector organisation?
- As part of the efforts aimed at establishing the necessary leadership and organisational culture how can clear roles, decision making on rights and accountability across the system be established?
- How can clear accountability for data governance within agencies be established?

4.3.3. Investment in data analytical tools

Alongside the skills and capabilities required by public sector workers, the necessary tools and equipment needs to be available to them (OECD, 2019[55]). Some of the tools that will help to extract the value from data will be expensive in their own right, deterring governments from making the necessary investment whilst others will require specific, targeted training. Such investment is not comparable with traditional data gathering methods such as large-scale surveying, and as a result governments may need to develop a different methodology for performing a cost-benefit analysis that will support any existing business case processes (OECD, 2019[55]).

Nevertheless, it is important to recognise that technology is not a panacea in terms of DDPS. The effective use of software that allows for the gathering of data analytics or the manipulation of existing data sources will require governments to respond to the challenges detailed in this section, particularly those related to the provision of useful data, the easy interoperability of data to enable reuse, and considerations related to legitimacy and public trust.

4.3.4. The role of the private sector in providing human and technical resources

The breadth and scope of the work performed by the public sector and the scale of the challenge it faces in responding to the opportunities of digital transformation mean that governments face a shortfall in the availability of capacity and capability. As many governments have limited knowledge, tools and infrastructure for handling data, those wishing to rapidly embrace a more data-driven approach to their activities may find that the private sector can offer an effective route to having the necessary resources at their disposal. IT infrastructure providers and data analytic service providers might play an important role in providing the necessary capacities to build a data-driven public sector.

However, where countries do decide to outsource the provision of services and the collection of data as a shortcut to acquiring the necessary skills and capabilities there is a

risk that the public sector will lack the necessary knowledge to make good commissioning decisions. Building on the experience of OECD countries, the Working Party of Senior Digital Government Officials (E-Leaders) has recognised this challenge and begun development of an ICT Commissioning Playbook (Box 4). Amongst other issues this highlights how critically important it is for governments to ensure contractual arrangements with third parties consider data control and intellectual property. When using private sector suppliers, governments should be explicit about the ownership of government data, including that created through the operation of a service, in order that it can be readily moved to a new supplier or deliver value to governments in its use and reuse.

Box 4. The ICT Commissioning Playbook

Currently in an Alpha phase of the development the ICT Commissioning Playbook has been produced by the Working Party of Senior Digital Government Officials (E-Leaders) Thematic Group on ICT Commissioning.

It focuses on ICT procurement reform and its part in the wider digital transformation of the public sector in countries around the world. Its goal is to show how traditional procurement can evolve towards agile procurement.

The playbook is for procurement professionals in the public sector: at all levels of government and its supporting agencies. The play presents outline ways to overcome common problems, alongside case studies that demonstrate real challenges and successes. These are based on the experiences of what works, and what does not in the United Kingdom alongside contributions from Australia, Canada, Chile, Finland, Mexico, New Zealand, Portugal, Uruguay, and the United States.

The playbook is designed to help people:

- learn from the procurement experiences of countries around the world
- use the lessons to begin to transform procurement in their own area of work
- provide feedback on the experiences of their own government or organisation
- share findings and collaborate with procurement professionals of all levels

Source: The ICT Commissioning Playbook Alpha https://playbook-ict-procurement.herokuapp.com/

Decisions about the extent to which governments might outsource the delivery of data-driven activity will often be driven by the budgetary and legal context in which governments are operating. These decisions should also be informed by the 2008 OECD Council Recommendation of the Council on the Protection of Critical Information Infrastructures[19].

4.3.5. Shared government platforms and data

A number of OECD governments have opted to develop common platforms as a way to share resources and ensure a more efficient use of technologies government-wide that avoid the duplication of efforts. For instance, the United Kingdom reported saving GBP 105

[19] The OECD has initiated a review procedure for this Recommendation, see DSTI/ICCP/REG(2016)2/REV1.

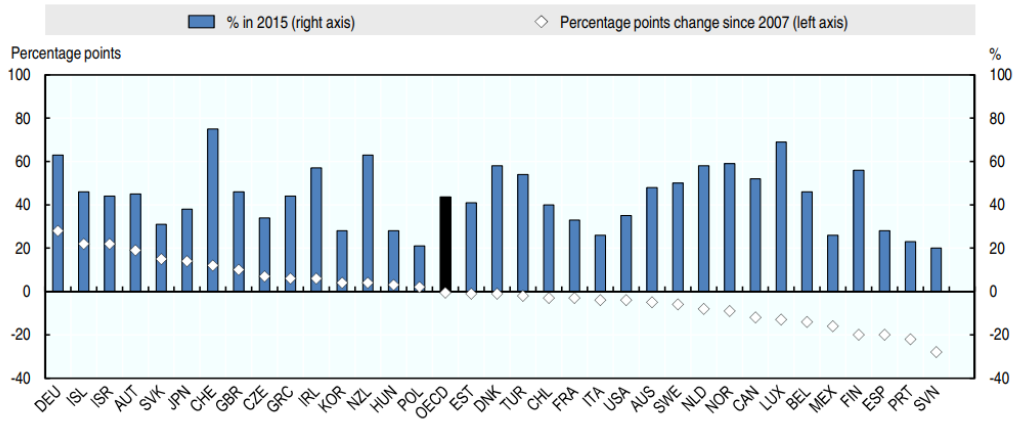
million in the 2014-15 financial year after eliminating multiple agency websites and creating the single government domain GOV.UK. This subsequently inspired the 'Government as a Platform' programme of work looking at providing reusable technology and resources to meet common needs in areas such as data, payments and notifications as well as incorporating existing activity related to identity.

The adoption of a whole-of-government approach has benefits in enabling the consolidation of skill, favouring interoperability and avoidance of costly duplication. Importantly from a data-driven point of view it provides a mechanism to build a more coherent view of how services are performing across the public sector. For example, in the case of payments, a disaggregated approach would mean data on the volume of payments, the preference for particular payment methods and the amounts themselves would be held locally in individual contracts and technology solutions. With a common payment platform government, it is able to generate an aggregated view of money being paid into government and can tailor its policy and procurement efforts accordingly.

4.4. Legitimacy and public trust

A final set of challenges facing governments wishing to maximise the role of data in their design of policy, delivery of services and efficiency of operations involves the legitimacy and trustworthiness of data-driven efforts in the eyes of the public. Figure shows that trust in governments is declining and DDPS approaches must ensure that they approach the use, and reuse, of data in ways that reinforce public trust and earn public support.

Figure 7. Confidence in national government in 2015 and its change since 2007



Source: OECD Trust Database, Gallup World Poll 2015 (OECD, 2017[11])

As such, governments must ensure that the management of data and its use is handled in a transparent, responsible and secure way that is proportionate to the expected benefits for society. These efforts create a legitimacy that is more than adherence to laws and other legal instruments. Legitimate government actions are justifiable according to legal, ethical and social norms, which can be interconnected and are all important to ensure public trust. Even in those countries where public trust maintains high levels, efforts should be made to ensure that the use and reuse of data does nothing to jeopardise this trust, particularly in the use of personal data.

This section will consider six areas which impact on legitimacy and public trust in efforts to embrace DDPS. They are:

- Ethical and accountable use of government data
- The imperative to act
- Protecting the privacy of citizens
- Transparency and establishing rights to data
- Managing risks to government data

4.4.1. Ethical and accountable use of government data

As more countries explore how they might successfully implement machine learning and other applications of AI, one important aspect to consider should be on ensuring access to high quality and useful data. However, it is equally important that questions of ethics and accountability are also considered.

Data analysts are not formal decision makers when it comes to the design and implementation of policy but their role in exploring problems and solutions is to provide suitable information to those with decision making power to make policy. In adopting a DDPS approach these analysts may find that the results of data analysis are in conflict with the politically motivated course of action in a given policy area. It is essential for governments to create the conditions in which all parties, whether public servant, politician, or citizen, have confidence that the use of data is serving the public interest

One of the considerations in doing this is being clear about what ethical use of data looks like. An important model for this is proposed in the OECD Recommendation of the Council on Public Integrity (OECD, 2017[56]) which defines public integrity as

> *the consistent alignment of, and adherence to, shared ethical values, principles and norms for upholding and prioritising the public interest over private interests in the public sector.*

But there are equally important considerations of privacy and other freedom rights to be considered in defining the ethical treatment of data in its applications in government. A helpful example of how a country has created the conditions in which its data practitioners work is the United Kingdom's development of a Data Ethics Framework within the Department for Digital, Culture, Media and Sport (GOV.UK, 2018[57])**.**. Endorsed at ministerial level this code of conduct is developed by the community of practice and complements other aspects of the UK's standard-based approach to policy making and service delivery to create a clear set of expectations.

One of the reasons for introducing AI is to help complement more traditional policy making and the decision-making activities. Machine learning algorithms can do more than provide information to decision-makers. In some cases, they have the potential to enhance and foster data-based decisions. A well-known example is the automation of the stock market, where algorithms predict developments in the financial market and consequently buy or sell on behalf of their clients. Approximately seventy percent of all financial market transactions are now made by automated trading algorithms (Helbing, 2015[19]).Algorithms are not only able to augment humans' intellectual capacities, but also physical capacities (OECD, 2015[16]). Indeed, it is not just in the virtual domain that such decisions are being made, the increasing presence of sensors and IoT devices means the physical world is producing data that lead to decisions. The emergence of data-based traffic management in

'smart cities' and 'smart grids' for automated electricity storage and distribution are striking examples (Tuballa and Abundo, 2016[58]).

When it comes to the decisions that are made by government they can deeply affect the lives of citizens or have long-term implications for their quality of life. As a result it is important to consider the extent to which such approaches should be applied. As digital transformation becomes the norm it will be increasingly hard to find policy domains that are exempt from the cultures, practices, processes and technologies of the internet-era. Full or partial automation of decision-making in these areas is not a question of preventing machines from supplanting humans but instead one of governing the use of data and associated tools by those involved in their design, collection and processing.

Humans normally decide where a data-driven approach to policy will be applied. Humans set the priorities for the outcomes they wish to see, and will have made the decisions recorded in whatever data is used to train any AI that is deployed. The computers performing the analysis do not decide by themselves the topics on which they will provide predictions or policy suggestions. The decisions which computers make are done on the basis of the data to which they have access. Therefore, ethical considerations are not only of relevance when interpreting the data, but already when selecting, gathering and generating data.

Governments need to be clear about who has control of data within a DDPS and work to ensure that any work is being done within ethical boundaries and according to democratic principles. The transparency around the use of algorithms is one way in which governments can attempt to build trust and accountability in data-driven decision making. The French Digital Economy Law (*Loi Lemaire*) specifies that algorithms used in public decision-making should be transparent to the public (except where they are involved in detecting fraud).

Whilst it may help to publish the algorithms that are used to form judgments, the most powerful demonstration in support of trusting AI are the outcomes of that judgment process. This requires government to be open about how machine learning algorithms are trained, the process by which final judgment is reached on their output, how an acceptable margin of error was defined for data-driven decisions, and a clear rationale for how any data-driven decisions compare to the standards expected of human-made decisions.

The OECD's Working Paper on the State of the Art on the use of Emerging Technologies in the Public Sector (Ubaldi et al., 2019[15]) discusses how governments are augmenting human capacity with technology. This is particularly present in those cases where machine learning activities using neural networks and algorithms begin to take over the responsibility for making decisions. As this happens, it becomes increasingly important to ensure citizens can establish that those decisions were made in a fair, democratic and inclusive way. One approach to this is to make algorithms openly available as specified by the French Digital Economy Law (*Loi Lemaire*). One important fact is the level of trust citizens have in the broader environment in which automated decisions are made and therefore governments need to find ways to communicate a consistent message that automated decision making isn't threatening. Nevertheless, this raises several questions.

- How can organisations explain the process of an automated decision in a way that people understand it?
- How can that process then be extrapolated to be assessed against some criteria of fairness? And who establishes that criteria?

- Irrespective of whether or not a decision can be explained, how can an organisation demonstrate that they are fair?
- How should such decisions be regulated? And investigated if something goes wrong?

Technology itself can perhaps be part of the solution to the opaqueness of AI-based decision-making. Contrary to traditional AI, Explainable Artificial Intelligence (XAI) attempts to provide the reasoning behind the performed evaluation as part of the output of the analysis. The XAI would not only state with a particular certainty that a person is likely to develop disease X, be a good fit for an analyst job in Ministry Y or commit social benefit fraud, but which combination of (risk) factors has led the AI to come to this conclusion.

The aim of XAI is to enable two-way communication between the machine-learning algorithm and the human decision-maker/supervisor using the information produced by the AI. If traditional AI fulfills the role of an expert giving advice without explaining the underlying reasons, XAI acts as an expert providing clarity about the parameters leading up to a certain advice with an openness to implement changes to those parameters if necessary. This would allow human decision-makers to ensure that a picture of a frog is not mistaken for that of a pool table (Ribeiro, Singh and Guestrin, 2016[59]).

XAI would not only facilitate government officials to judge the value of the AI-expert advice but would also prevent decision-makers of lesser integrity from manipulating the algorithmic black box to their advantage or blaming 'the computer' if things did not turn out as expected. Crucially, the transparency offered by XAI would allow stakeholders to assess the underlying assumptions behind these data-driven decisions and consequently have the power to hold decision-makers accountable.

Whether or not a technological approach like XAI provides the solution, it is necessary for governments to understand how they might audit their decision making processes and analyse the outcomes which affect citizens lives in order to ensure their data-driven decisions deliver on their hopes and expectations.

The accountability of services and policies is an important issue in building legitimacy and public trust in DDPS approaches. This is not only from the perspective of ensuring that citizens can understand the basis on which a decision affecting their life was taken but from a practical perspective too. As governments seek ways to increase their transparency in general, they will need to establish effective mechanisms for explaining the role which algorithms have played. This may lead to governments using rather straight-forward and simple algorithms rather than complex machine-learning, neural network, based algorithms which are difficult to explain.

4.4.2. Protecting the privacy of citizens

DDPS approaches can help governments increase the effectiveness and quality of the services they offer to citizens and businesses. As the economy is stimulated and well-being enhanced, these interventions can also increase levels of trust in government. However, should that data which is the responsibility of government be misused then that trust will be threated. Consequently, another challenge facing governments wishing to maximise the opportunities of data is to consider how they might approach its use in ways that protect the privacy of citizens in a holistic fashion, considering bodily privacy, and intellectual privacy.

A considerable amount of the data gathered, processed and controlled by the public sector is personally identifiable. Whilst this allows for governments to design and develop more proactive approaches to service delivery, it may also be viewed as being invasive to a person's personal life. Governments can distribute personalised invitations to complete tax returns and use the aggregation of individual mobility data to improve traffic management. These are possibly benign interventions but consider the implications of those fearing government over-reach if public sector organisations use personally identifiable information from several sources to begin refusing access to particular services or monitoring behaviour in order to confirm, or deny, eligibility of a given benefit.

The perceived misuse of government personal data may come from the failure to secure it effectively or selling data to private companies. It may also come from well-meaning efforts to genuinely improve outcomes by partnering with third parties. In the United Kingdom, a collaboration between Deepmind and the Moorfield Eye Hospital saw the use of patient data to train an Artificial Intelligence, *Streams*, for supporting nursing staff. Whilst this arrangement brought benefits in patient care, the UK's Information Commissioner ruled that the UK hospital trust had broken UK privacy law in failing to tell patients about the way that their data was being used (Denham (UK Information Commissioner), 2017[60])

The circumstances under which such activities are legal and justifiable is the subject of privacy protection regimes. The OECD has an established role in promoting respect for privacy as a fundamental value and condition for the free flow of personal data across borders. The OECD Privacy Framework (OECD, 2013[61]) recommends several principles for the handling of personal data:

- Collection limitation principle
- Data quality principle
- Purpose specification principle
- Use limitation principle
- Security safeguards principle
- Openness principle
- Individual participation principle
- Accountability principle

Such guidance can help governments identify how to safeguard their citizens' data and the legal, technical, or informational, measures that are necessary to implement to support that safeguarding. The questions that these principles raise are relevant not only for citizens as clients of public institutions, but also for public servants as they go about their jobs.

Legal norms at the international, national and regional/state level can provide an helpful mechanism in understanding the restrictions being placed on what countries can do with data (particularly around reuse) and the obligations placed on their activity in how they handle that data (for example in providing access to information). The European Union's General Data Protection Regulation (GDPR) came into force in May 2018 and replaced the Data Protection Directive 95/46/EC not only as a means of harmonising data privacy laws across Europe but providing a new baseline for protecting and empowering EU citizens in accessing their own data.

However, as citizens enjoy and expect increased protections over their personal data, governments are faced with the challenge of preserving some of the economic interests that could be derived from the use of that data. This raise several important considerations about the circumstances under which a government should be able to access and use personal data.

Under what circumstances, for example, should governments be able to access business data, especially if it relates to consumer habits and activity? Or when can government delegate the responsibility for the collection and production of citizens' data to the private sector?

Equally, there have been challenges for government in how best to share data with businesses (G2B). This is especially relevant where service delivery has been outsourced to a private company and where the quality of these services could be enhanced with access to existing government data. Within the European context, this is being addressed through the reform of EU Data Protection in order to facilitate the European Digital Single Market[20].

Several fundamental rights of citizens are at risk as governments begin to explore (predictive) big data analytics, profiling and data-driven classification activities, especially in the domain of security policies. These risks include breaches of privacy, discrimination and stigmatisation as well as chilling effects on the freedom of speech (OECD, 2015[16]).

There are further issues that may impact on the lives of citizens. Data analytics may result in greater efficiencies, but may also limit an individual's ability to modify path-dependent trajectories and escape socio-economic lock-ins (OECD, 2016[62]). Data and analytical programmes could serve to manipulate people, distort their perception of reality and influence their choices (Glancy, 2012[63]; Helbing, 2015[64]). As a result, individual autonomy, free thinking and free will are challenged, potentially undermining the foundations of modern democratic societies.

Furthermore, biases in the data itself, human biases and the unethical use of data analytics pose a fundamental risk to freedom rights. Consequently, those working with data must be focused on how they can detect these effects, measure their impact and provide any necessary correction. Governments need to consider how procedures might be designed to ensure the legitimate treatment of conflicting values such that any correction does not exacerbate the underlying issue. Therefore, governments must consider legal, ethical and social norms in their approach to such situations. The complexity of this issue is demonstrated when looking at privacy regimes worldwide, as for example the expected effectiveness and feasibility of the article in the new EU General Data Protection Regulation (GDPR) covering non-discrimination regarding algorithmic profiling (Goodman and Flaxman, 2016[65]).

4.4.3. Transparency and establishing rights to data

As more people become concerned about how their data are being used and legislation enshrines protections in law, there will be an increased expectation for governments to be transparent about how they handle data, and for citizens and businesses to expect to have greater control over how it is used throughout the Government Data Value Cycle (Figure 2). Discussions around citizens' rights? How much control should they have on access,

[20] For more information see http://ec.europa.eu/justice/data-protection/reform/index_en.htm (accessed 25 July 2016).

modification and use of personal data held by government? Should existing rights be strengthened or extended?

When data are being **collected** or **gathered** the challenge of being transparent over where data have come from can appear to be straightforward but, given concerns over a surveillance state it may be more controversial than is first expected. Should an expectation of transparency be created then any time where information is kept closed will likely have repercussions in terms of trust (OECD, 2015[16]).

When data are being **processed** and **stored**, should there be transparency over the databases and the arrangements around how they are managed? Data 'at rest' in this stage of the cycle are often the most susceptible to digital security threats. Consequently, does transparency here increase, or reduce, the possibility of a data breach? Whilst governments should be clear about how long they will retain data and the circumstances under which they remove it, what rights does a citizen have to be forgotten? Whilst such a right is codified within GDPR, is it appropriate for that to extend to the government who provides you with goods and services? (OECD, 2015[16]).

The **sharing** of data is important for governments to be able to provide transformative services that avoid citizens or businesses having to provide the same data to multiple sources. Providing transparency when data is shared, or **reused**, is being explored in various ways but could emulate a model from the Netherlands where citizens can carry out a "personal data registry scan" (via https://www.wiekrijgtmijngegevens.nl (whogetsmydata)) which shows the organisations with the ability to retrieve data from the personal records base register.

Finally, when data are being used for **decision making** how transparent should a government be. It is unlikely to be a single approach for all data within government, do some situations and topics require different approaches? If a decision has been carried out following algorithmic assessment then if a citizen has the right to an explanation, what does that mean in practice and how much can the explanation of how an algorithm reaches a decision be understood? As seen in previous sections, the French Digital Economy Law (*Loi Lemaire*) is a good example. The French government has made algorithms used for public decision making in the openly available.

There are several arguments to be made in favour of data transparency. Being transparent about data characteristics and data use isn't only a matter of enhancing accountability in terms of how taxpayer money is spent but of empowering people in their democratic rights and ability to increase their own well-being.

Nevertheless, being transparent does not always have a positive effect on public trust in government (Grimmelikhuijsen et al., 2013[66]). Moreover, transparency on the detailed functioning of data analytics may open up the door for manipulation of data and consequently influence the output of analytics (Domingos, 2015[22]). At the same time, with efforts to try and open up the policy making process to public scrutiny the introduction of data-driven approaches could in fact deliver more opacity and undermine government transparency (Lammerant and Hert, 2016[67]).

4.4.4. Managing digital security risks

The digital revolution is transforming not only the way in which governments deliver services but the behaviours of criminal and military elements who may threaten the security of citizens through various digital security threats. In recent years several governments have been impacted through the crippling of computer systems due to malware like

WannaCry and NotPetya or the breach of valuable data sets, such as voter data, citizenship data and personnel data. In order to respond to such threats, and safeguard against them there must be clear governance guidelines and proactive efforts by government to keep their citizens safe. Citizens need to have the confidence everything is being done to keep their information secure.

The OECD Recommendation on Digital Security Risk Management (OECD, 2015[68]) provides guidance on the economic and social prosperity dimensions of digital security risk. It highlights the importance for governments to effectively and efficiently manage threats, vulnerabilities and incidents related to data security (OECD, 2015[68]).

One way in which government can handle this risk is in the expectations it sets for its staff whether employed directly, or working through contracting suppliers. Policies regarding data storage, inventories, encryption, access and identity management are crucial to managing any digital security risks but must be enacted within organisations rather than being something that can be set across government. Furthermore, it is the responsibility of each service to consider how they monitor and secure their production environments to keep personal data safe from falling into the wrong hands or being manipulated and other data from being compromised (OECD, 2015[68]).

In response to these threats several OECD countries including Portugal and the United Kingdom, have established dedicated organisations to handle questions of cybersecurity both in keeping the country safe and helping government departments to implement effective policies.

5. Enabling frameworks

This final chapter considers the coherent strategic approach that is needed for governments to effectively implement DDPS. This covers digital government, open government data, public sector innovation, public sector productivity, human resource management, financial and budgetary management, risk management and the way data is used across sectors. Finally, it defines the strategic enablers that governments need in order to maximise DDPS activities.

5.1. A coherent strategic approach to data governance across the public sector

Developing a better understanding of how data can transform the entire public sector leads to a recognition of how a **whole-of-government strategy for public sector data governance** can help to best achieve this transformation.

The OECD's model for data governance (see Figure 1) provides a framework for addressing the several elements required for successfully implementing the DDPS. These are not technical tasks addressed by recruiting data professionals but a strategic set of activities, capacities, roles and instruments that aims to transform data into an asset for business intelligence. Although some of these elements may already be in place, the lack of coordination and coherence in delivering value, instead reflect siloed attitudes towards how data are managed and shared.

A coherent strategic approach to data governance is a prerequisite to realise a data-driven transformation of the public sector. It is indeed a critical factor to the adoption of a "data as a strategic asset", perspective which underpins a successful use of data to support improved policymaking and service delivery. It would state the relevant enabling frameworks (e.g. laws, budget and governance structure) and provide simple guidance for a robust management of data as a key asset throughout its entire lifecycle. Countries would be able to monitor and evaluate new and existing data-driven initiatives, thereby strengthening their value proposition and would be able to identify the need for new data and those datasets which are redundant.

It would support better use, flow and sharing of public sector data within and across government, and across borders, e.g. by improving standardisation and facilitating data cataloguing, discoverability, and traceability. It would also support prioritisation of efforts in the production of quality data in line with standards and guidelines, possibly supporting non-institutional stakeholders (for example in the private sector) to collect and/or produce certain data on behalf of the public sector. Such a strategy would help to better identify and meet the needs of citizens while increasing efficiencies for both service users and public servants.

A strategic approach to data governance, inclusive of developing the enabling frameworks, is essential for removing siloed approaches to data management - that unless justified by a particular and legitimate purpose (e.g. security/privacy) - represent important obstacles,and helping governments move beyond the hype surrounding data-driven measures, and the volatility of incidental or ad-hoc data-based decisions, towards more solid and sustainable

data-driven public sectors. Such an approach is needed to recognise data as an essential ingredient, along with people and money, of public sector performance.

Box 5. The US Federal Data Strategy

The US President's Management Agenda (PMA), released in March 2018, focuses on agencies' mission, service and stewardship and presents a vision to modernize the Federal Government. The Agenda outlines the importance to create a Federal Data Strategy to improve the use and management of Federal government data. It is expected to include four key components:

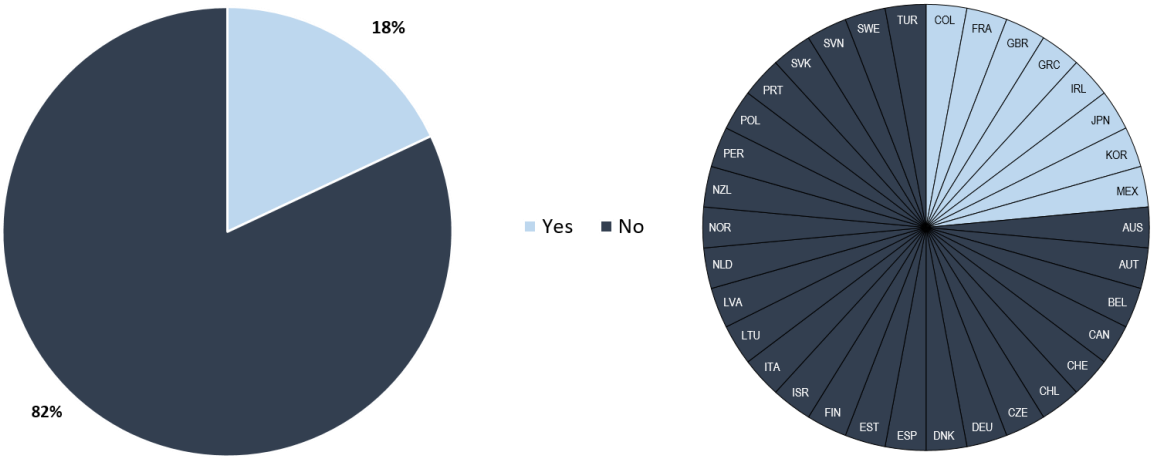
1. *Enterprise Data Governance.* Set priorities for managing government data as a strategic asset, including establishing data policies, specifying roles and responsibilities for data privacy, security, and confidentiality protection, and monitoring compliance with standards and policies throughout the information lifecycle.
2. *Access, Use, and Augmentation.* Develop policies and procedures that enable stakeholders to effectively and efficiently access and use data assets.
3. *Decision-Making and Accountability*. Improve the use of data assets for decision-making and accountability for the Federal Government, including both internal and external uses.
4. *Commercialization, Innovation, and Public Use.* Facilitate the use of Federal Government data assets by external stakeholders at the forefront of making government data accessible and useful through commercial ventures, innovation, or for other public uses.

Source: Roundtable on Leveraging Data as a Strategic Asset (The Centre for Open Data Enterprise, 2018[69])

Moreover, the development of a joint strategic enterprise-wide approach to data governance and the design of data-driven services and organisations, can help to prevent the fragmentation of technical systems, build a data culture among civil servants and deliver sustainable, flexible policies and services designed in collaboration with citizens, businesses and civil society.

A common strategy is a critical success factor in supporting clearer data governance that provides the basis for identified roles and responsibilities and government-wide leadership for the data-driven agenda. In 2014, only a minority of countries reported having a chief data officer at the central/federal government level, whereas in 2016 the number of OECD countries that has created this position has slightly increased.

Figure 8. Existence of a Chief Data Officer for central government in 2014 compared to 2016



Note: Data for 31 OECD countries + Colombia and Lithuania
Source: OECD Survey on Open Government Data (2014)

Note: Data for 33 OECD countries + 1 OECD partner and other economies (Peru)
Source: OECD Survey on Open Government Data 3.0 (2017)

5.2. Aligning cross-cutting strategies

The overall challenge of rethinking government strategies and frameworks for data governance necessarily entails a reflection on the broader strategies and policies in a number of other policy areas where data already plays an important role. This is essential to guarantee policy coherence. How can governments move from fragmented, silo-structured data management frameworks to a single coherent cross-government vision supporting a strategic use of data across the whole public sector that impacts on all policy areas?

5.2.1. Digital government

Strategies for data management on the one hand and digital government on the other must be considered in partnership with one another. Data are becoming indispensable for a growing number of digital government activities, such as personalised digital services, social media strategies and digitally supported administrative systems. At the same time, technology is needed across the entire data value-chain. As a result, a digital government strategy must focus on creating the necessary conditions for stakeholder engagement, governance and implementation capacities that ensure digital technologies and approaches create value. Data management strategies, policies and governance should be part of an overall vision and align with digital government policies to ensure coherent objectives and practices across government.

5.2.2. Open Government Data

A number of these issues are also challenges for Open Government Data policies. Since open government data, next to undisclosed government data and data produced or reused by citizens and businesses, are an important resource for the data-driven public sector. Therefore, it is essential that strategies, which aim to develop a data-driven public sector

align with OGD strategies aimed at promoting OGD reuse within the public sector, data-driven collaborations with citizens and development of joint solutions.

Furthermore, improving the discoverability of data and cooperation across public sector organisations, for example by developing and providing access to data inventories, is a common challenge for open data policies and data management strategies. Finland announced its open data policy for 2015-2020, which focuses on 'open data as a basis for the digital economy, public services and decision-making', thereby becoming an integral part of the Finnish digitalisation strategy (Finnish Ministry of Finance, 2015[70]).

5.2.3. Public sector innovation

In order for public sector organisations to find new approaches to create public value for individuals and society, they must be able to identify opportunities for innovation whilst applying the lessons of their successes and failures. This requires governments who are interested in supporting innovation to create environments in which the barriers to experimentation have been removed. As a result, governments must prioritise data sharing and the creative use of data.

5.2.4. Public sector productivity

Improving public sector productivity necessarily entails measuring public sector inputs, outputs and outcomes, understanding it and developing a strategy for the future. What provisions should governments take to enable the use of data and analytics in support of these efforts?

5.2.5. Human resource management

HRM strategies and data and information strategies are mutually enabling. Sound data management is needed to provide human resource managers with the necessary data input for insight and oversight to address organisational health issues. At the same time, it is crucial to integrate data skills management into human resource management to ensure a capable data-driven public sector.

5.2.6. Financial and budgetary management

The ways in which governments plan, and then execute, the spending of taxpayer money is heavily scrutinised within government and by media and civil society voices outside it. Beyond the way in which data can be used to increase accountability and build trust, how can countries use predictive modelling and ongoing data analysis to better plan expenditure, stimulate policy making and respond to forthcoming crises in ways that minimise their impact?

5.2.7. Risk management

Data management is of crucial importance at every stage of the risk management cycle: in translating unforeseeable events into calculable risks, stakeholders are empowered to take action in anticipation of new risks and in reaction to developing crises. At the same time, efforts to leverage data to address inefficiencies and unethical behaviour in public institutions may benefit from a risk management approach.

5.2.8. Sector-specific strategies and policies

In more and more policy fields, such as infrastructure, national security and law enforcement, environment, tax policy and urban policy, the topic of better data usage is on the agenda. Fortunately, this is matched by an awareness that sector-specific data strategies cannot be developed in isolation.

Data for the planning, delivery and management of transport services and infrastructure is collected, stored and exploited by a diverse set of actors that extends well beyond the field of transport (International Transport Forum, 2016[71]). Similarly, strategies for exchanging information to address international tax evasion are more effective when aligned with overall data management strategies that can ensure governments deliver on their international commitments in this area (OECD, 2016[72]). The important role of data for the development of a National Urban Policy was underlined at the UN conference Habitat III, which set out the urban agenda for the next 20 years. Building on 'Smart City' initiatives, a central recommendation is to ground national urban policy in the most current and comprehensive qualitative and quantitative data available (United Nations General Assembly Habitat III, 2016[73]). Additionally, a government-wide data management strategy is a key enabler for data-related ambitions in specific policy areas and their interconnectedness, possibly bringing about systemic changes. For example, analysing transport data in conjunction with environmental data can help reshape both urban traffic management, housing policies and environment protection systems.

Therefore, it is essential that governments adopt a strategic approach to the governance of data and delivery (Figure 1), which transcends government silos.

5.3. Summarising the strategic key enablers for a data-driven public sector

In addition to the definition and adoption of common strategy defining a shared vision, if governments want public organisations, policies and services, and ultimately society, to benefit from the use of data in a sustainable manner, while safeguarding public trust, it is essential to develop and put in place the enablers for data to flow through the data value cycle.

Box 6. The Italian Data & Analytics Framework (DAF)

Enabling the public sector to make the most of data to the benefits of organisations and citizens

Italy plans to introduce the National Digital Data Platform (PDND) to overcome silos within public sector organisations being able to leverage data as a common good to be explored and mined to extract value. The idea is to create a new interface enabling public sector organisations to communicate and share data and APIs in a free and open way, and thus allow for the creation of new services and data applications that respond to citizens' needs. The Italian Code for Digital Administration (CAD) acknowledges the Data and Analytics Framework (DAF) and introduces the PDND. The PDND is the institutional reference for the DAF project introduced in the Italian Three-Year Plan for ICT in the Public Administration.

The DAF is expected to improve and simplify the data interoperability and exchange between public sector organisations, promoting and improving the management and use of Open Data, optimizing activities of analysis and knowledge generation. The idea is to open the world of Public Administration to the benefits offered by modern Big Data management and analysis platforms along three main lines:

- Enhancing the value of government data as assets through the preparation and use of analytical tools designed to synthesise knowledge for decision makers, and the dissemination of information to citizens and businesses.
- Optimizing data exchange between public sector organisations and Open Data deployment, minimizing transaction costs for data access and use.
- Facilitating data analysis and data management by data scientist teams within the public sector to improve knowledge of phenomena described by the data and develop "intelligent" applications and promote scientific research activities on application themes of interest to the public sector

Source: Italian Digital Transformation Team (https://teamdigitale.governo.it/en/projects/daf.htm)

These enablers not only encompass the technical and operational capacities (implementation framework), but also the legal, ethical and social provisions (normative framework). Both levers for public trust, policy competence and good governance, benefit from having a sound normative framework for data governance.

- **Data-driven policy competence** - guidance and direction from legal framework, eliminating legal barriers to achieving policy results and ensuring social support for data-driven activities;
- **Good data governance** - data-driven governments are more likely to be perceived as fair and responsible if they demonstrate having given thought to translating principles such as integrity and openness into their data governance strategy.

The implementation framework should always be developed in close conjunction with the normative framework to ensure that the desired values and principles find their place in the public sector data strategy or policy and its implementation. The schematic overview below of key enablers for each stage of the data value cycle can be used to map countries' data-driven landscapes.

Figure 9. Enabling frameworks for a data-driven public sector

		Collecting, generating and processing data	Sharing data	Generating information from use/re-use of data	Decision-making with data
Normative framework	Creating the conditions for the treatment of data to be **legally acceptable**	*e.g. data retention laws*	*e.g. open data law; Transparency law; Copyright law*	*e.g. legal restrictions on using data from other organisations*	*e.g. algorithmic transparency law*
	Creating the conditions for the treatment of data to be **ethically acceptable**	*e.g. principles on the outsourcing of storing personal data of citizens*	*e.g. open by default principle*	*e.g. a data ethics framework regulating the purposes for which data is used*	*e.g. identifiable accountability for automated decisions*
	Creating the conditions for the treatment of data to be **socially acceptable**	*e.g. privacy perception surveys; public debate*	*e.g. expectations around handling that puts control in the hands of the citizen*	*e.g. the 'once only' principle avoids duplicative effort and simplifies services*	*e.g. communicating how decisions are made with data*
Implementation framework	Building on the normative framework, actions to take in the area of **governance**	*e.g. appoint a Chief Data Officer*	*e.g. provide Open Government Data support teams; guides on personal data processing; data sharing agreements; standards for data interoperability and metadata;*	*e.g. produce a data strategy which maps the use, re-use and gaps in government data for policies and services*	*e.g. create a community of practice for data-driven decision makers across the public sector*
	Building on the normative framework, actions to take in the area of **technology**	*e.g. unblock the procurement or use of cloud solutions for data storage*	*e.g. an interoperability platform; standards for data interoperability and metadata; canonical sources of data*	*e.g. machine learning algorithms; data visualisation tools*	*e.g. publicly visible dashboards of performance*
	Building on the normative framework, actions to take in the area of **culture**	*e.g. culture change activities and behaviour modelling from senior leadership to encourage, champion and support a data-driven approach across each area; strategic activities and communities of practice designed to overcome stove-piped and silo based practices*			
	Building on the normative framework, actions to take in the area of **skills**	*e.g. individual and corporate training on an ad hoc or systematic basis to broaden data skills across policy, delivery and operational disciplines targeting each of these areas*			

6. Conclusion

Strategic leveraging of data management and use (e.g. data analytics) has the potential to boost public sector intelligence, thereby supporting public sector reform towards more inclusive and sustainable governance. In a DDPS, organisations can work more efficiently, become more productive and have the capacities to transform the design and delivery of public policies and services in ways that improve social goods and impact societal well-being. In turn, this can have a positive impact on public trust.

In order to realise this potential, it is critical for governments to explore and understand the innovative and transformative value of data and create the necessary institutional conditions to leverage this value.

In line with the above, this Working Paper is intended to shift the focus from the hype to the enabling factors needed to mainstream DDPS practices across the public sector and thus realise the potential of using data as strategic asset to improve public governance. The paper discussed some of the main opportunities created by DDPS for governments to be better prepared for the future, by being able to predict trends and patterns to mitigate emerging risks and respond to developing crises whilst creating the conditions for robust, evidence-based policy making. It then presented the opportunities for DDPS to adjust service delivery to better respond to user needs, and explored those concerning performance management, in terms of improved public sector productivity and the ongoing strengthening of how governments evaluate policies and impact.

To complement the opportunities, the paper also identified several challenges which governments may find when they come to implement a DDPS. These include, among others, limited availability and access to useful and quality data, inadequate data governance, organisational resistance, barriers to interoperability, and concerns on skills and capabilities needed to make the most of data. The paper stressed the importance for governments to take these issues into account and adequately address them to enact the transition towards a data-driven public sector.

The paper argued also that in order for a DDPS approach to work, the overall data governance should support a broad whole-of-government data vision. This would enable to mainstream isolated efforts and practices and equip the public sector with the necessary governance mechanisms, technical means, skills and culture, underpinned by the requisite legal, ethical and social frameworks.

To support governments in this endeavour, the Working Paper sets out a framework for analysing existing data related practices, identifying policy implications for implementing a DDPS approach and the efforts required to maximise its benefits and provides reference to emerging practices across OECD member countries.

Nevertheless, the above also raised several questions about how countries can go about implementing some of these ideas and tackling the barriers they encounter. As such, this Working Paper paves the way to future work for the Secretariat in considering the comparative experiences of OECD countries. This exercise will discuss:

- creating the conditions for a DDPS – starting with the establishment of the necessary model for data governance to secure leadership, skills and technological capabilities;
- adopting ethical approaches to the management and use of government data (particularly citizen data) and the need to safeguard and explore emerging digital rights;
- developing business cases for DDPS by identifying potential gains, proposing mechanisms to achieve these gains and highlighting priority areas of investment.
- mapping the possible costs of inaction and inability to mainstream DDPS practices, for example in terms of financial loss (e.g. the inability to address data-savvy fraud) or exacerbating the trust deficit (e.g. the inability of responding to expectations of citizens).

References

Bulger, M., G. Taylor and R. Schroeder (2014), *Data-Driven Business Models: Challenges and Opportunities of Big Data*, https://www.oii.ox.ac.uk/archive/downloads/publications/nemode_business_models_for_bigdata_2014_oxford.pdf (accessed on 15 February 2019). [48]

De Feraudy, T. (2017), *Yes, 'govtech' can change the way cities function. New lessons in open government from Paris*, Citiscope. [30]

Deloitte (2016), *Global Human Capital Trends 2016*, http://www.deloitte.com/us/hctrends (accessed on 14 February 2019). [38]

Denham (UK Information Commissioner), E. (2017), *RFA0627721 – provision of patient data to DeepMind*, Information Commissioner's Office, https://ico.org.uk/media/2014353/undertaking-cover-letter-revised-04072017-to-first-person.pdf (accessed on 4 December 2018). [60]

Desouza, K. and B. Jacob (2014), "Big Data in the Public Sector: Lessons for Practitioners and Scholars", *Administration & Society*, Vol. 49/7, pp. 1043-1064, http://dx.doi.org/10.1177/0095399714555751. [23]

Domingos, P. (2015), *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, New York. [22]

Ehlers, R. (2018), *The Data Care Act of 2018 | Workplace Privacy, Data Management & Security Report*, Workplace Privacy, Data Management & Security Report, https://www.workplaceprivacyreport.com/2018/12/articles/data-security/the-data-care-act-of-2018/#page=1 (accessed on 26 February 2019). [26]

Erickson, J. (2014), *Algorithms Lecture 0: Introduction*, http://www.cs.uiuc.edu/~jeffe/teaching/algorithms/ (accessed on 19 April 2017). [18]

European Union (2009), *Ministerial Declaration on eGovernment*, https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/ministerial-declaration-on-egovernment-malmo.pdf. [53]

Finnish Ministry of Finance (2015), *Open Data Programme*, Finnish Ministry of Finance, Helsinki, https://vm.fi/documents/10623/360816/Avoimen+tiedon+tavoitteet+2015-2020/c7e9c09c-c492-4f04-ac52-2449b0c4973b/Avoimen+tiedon+tavoitteet+2015-2020.pdf. [70]

Gerson, D. (2015), *Lead-Engage-Perform: Enhancing Employee Engagement to Foster Innovation and Productivity in the Public Sector*, OECD, Paris, https://one.oecd.org/document/GOV/PGC/PEM(2015)2/en/pdf (accessed on 15 February 2019). [37]

Ghavami, P. (2015), *Big Data Governance: Modern Data Management Principles for Hadoop, NoSQL & Big Data Analytics*, Peter K. Ghavami. [4]

Glancy, D. (2012), "Privacy in Autonomous Vehicles", *Santa Clara Law Review*, Vol. 52/4, http://digitalcommons.law.scu.edu/lawreview. [63]

Goodman, B. and S. Flaxman (2016), "European Union regulations on algorithmic decision-making and a "right to explanation"", http://dx.doi.org/10.1609/aimag.v38i3.2741. [65]

GOV.UK (2018), *Data Ethics Framework*, GOV.UK, https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework (accessed on 5 December 2018). [57]

Government Accountability Office (2015), *A Framework for Managing Fraud Risks in Federal Programs*, Government Accountability Office, https://www.gao.gov/assets/680/671664.pdf (accessed on 15 February 2019). [45]

Griliopoulos, D. (2014), "Data exchange helps humanitarians act fast and effectively", *The Guardian*, https://www.theguardian.com/global-development/2014/dec/29/humanitarian-data-exchange-ebola-refugees. [5]

Grimmelikhuijsen, S. et al. (2013), "The Effect of Transparency on Trust in Government: A Cross-National Comparative Experiment", *Public Administration Review*, Vol. 73/4, pp. 575-586, http://dx.doi.org/10.1111/puar.12047. [66]

Grundke, R. et al. (2018), *Which skills for the digital era?: Returns to skills analysis*, OECD Publishing, https://dx.doi.org/10.1787/9a9479b5-en. [54]

Helbing, D. (2015), "Societal, Economic, Ethical and Legal Challenges of the Digital Revolution: From Big Data to Deep Learning, Artificial Intelligence, and Manipulative Technologies", http://arxiv.org/abs/1504.03751 (accessed on 17 February 2019). [64]

Helbing, D. (2015), *Thinking Ahead - Essays on Big Data, Digital Revolution, and Participatory Market Society*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-15078-9. [19]

Höchtl, J., P. Parycek and R. Schöllhammer (2016), "Big data in the policy cycle: Policy decision making in the digital era", *Journal of Organizational Computing and Electronic Commerce*, Vol. 26/1-2, pp. 147-169, http://dx.doi.org/10.1080/10919392.2015.1125187. [28]

International Transport Forum (2016), *Data-Driven Transport Policy*, International Transport Forum, https://www.itf-oecd.org/sites/default/files/docs/data-driven-transport-policy.pdf (accessed on 31 August 2018). [71]

Isson, J. and J. Harriott (eds.) (2016), *People Analytics in the Era of Big Data*, John Wiley & Sons, Inc., Hoboken, NJ, USA, http://dx.doi.org/10.1002/9781119083856. [36]

Janssen, R. (2015), *The Art of Audit. Eight remarkable government auditors on stage :*, http://www.oapen.org/record/619590. [44]

Klaassen, B. et al. (2017), "Usability Evaluations of a Wearable Inertial Sensing System and Quality of Movement Metrics for Stroke Survivors by Care Professionals", *Frontiers in Bioengineering and Biotechnology*, Vol. 5, p. 20, http://dx.doi.org/10.3389/fbioe.2017.00020. [50]

Ladley, J. (2012), *Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program*. [3]

Lammerant, H. and P. Hert (2016), *Predictive profiling and its legal limits: Effectiveness gone forever*, Amsterdam University Press/WRR, https://research.tilburguniversity.edu/en/publications/predictive-profiling-and-its-legal-limits-effectiveness-gone-fore (accessed on 18 February 2019). [67]

Lanier, J. (2010), *Who Owns the Future?*, Simon & Schuster. [24]

Lazer, D. et al. (2014), "The Parable of Google Flu: Traps in Big Data Analysis", *Science*, Vol. 343/6176, p. 1203, http://dx.doi.org/10.1126/science.1248506. [51]

Linturi, R. et al. (2014), *100 Opportunities for Finland and the world: Radical Technology Inquirer (RTI) for anticipation/evaluation of technological breakthroughs*, http://www.parliament.fi (accessed on 3 December 2017). [20]

Martin, E. and G. Begany (2016), "Opening government health data to the public: benefits, challenges, and lessons learned from early innovators", *Journal of the American Medical Informatics Association*, p. ocw076, http://dx.doi.org/10.1093/jamia/ocw076. [33]

Martin, K., A. Borah and R. Palmatier (2017), "Data Privacy: Effects on Customer and Firm Performance", *Journal of Marketing*, doi: 10.1509/jm.15.0497, pp. 36-58, http://dx.doi.org/10.1509/jm.15.0497. [52]

McAfee, A. and E. Brynjolfsson (2012), *Big Data: The Management Revolution*. [42]

McKinsey Global Institute (2013), *Game changers: Five opportunities for US growth and renewal*, http://www.mckinsey.com/mgi. (accessed on 14 February 2019). [41]

McKinsey Global Institute (2011), *Big data: The next frontier for innovation, competition, and productivity*, http://www.mckinsey.com/mgi. (accessed on 14 February 2019). [40]

Mickoleit, A. (2014), "Social Media Use by Governments: A Policy Primer to Discuss Trends, Identify Policy Opportunities and Guide Decision Makers", *OECD Working Papers on Public Governance*, No. 26, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jxrcmghmk0s-en. [47]

Morabito, V. (2015), *Big Data and Analytics*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-10665-6. [31]

OECD (2019), *Digital Government Review of Argentina*, OECD Publishing. [2]

OECD (2019), *Going Digital: Shaping Policies, Improving Lives*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264312012-en. [55]

OECD (2019), *Using digital technologies to improve the design and enforcement of public policies*, https://doi.org/10.1787/20716826. [34]

OECD (2018), *Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact*, OECD Digital Government Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264305847-en. [8]

OECD (2017), *Compendium of good practices on the use of open data for Anti-corruption*, http://www.oecd.org/gov/digital-government/g20-oecd-compendium.pdf (accessed on 14 February 2019). [43]

OECD (2017), *Digital Government Review of Norway: Boosting the Digital Transformation of the Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264279742-en. [10]

OECD (2017), *Recommendation of the Council on Public Integrity*, OECD, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0435 (accessed on 5 December 2018). [56]

OECD (2017), *Trust and Public Policy: How Better Governance Can Help Rebuild Public Trust*, OECD Public Governance Reviews, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264268920-en. [11]

OECD (2016), *OECD Science, Technology and Innovation Outlook 2016*, OECD Publishing, Paris, https://dx.doi.org/10.1787/sti_in_outlook-2016-en. [62]

OECD (2016), *Open Government Data Review of Mexico: Data Reuse for Public Sector Impact and Innovation*, OECD Digital Government Studies, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264259270-en. [13]

OECD (2016), *Technologies for Better Tax Administration: A Practical Guide for Revenue Bodies*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264256439-en. [72]

OECD (2016), "The Internet of Things: Seizing the Benefits and Addressing the Challenges", *OECD Digital Economy Papers*, No. 252, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jlwvzz8td0n-en. [39]

OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264229358-en. [16]

OECD (2015), *Digital Security Risk Management for Economic and Social Prosperity: OECD Recommendation and Companion Document*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264245471-en. [68]

OECD (2015), *OECD Public Governance Reviews: Estonia and Finland: Fostering Strategic Capacity across Governments and Digital Services across Borders*, OECD Public Governance Reviews, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264229334-en. [32]

OECD (2015), *Open Government Data Review of Poland: Unlocking the Value of Government Data*, OECD Digital Government Studies, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264241787-en. [12]

OECD (2013), *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188. [17]

OECD (2013), *The OECD Privacy Guidelines*, http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf. [61]

OECD (2003), *The e-Government Imperative*, OECD e-Government Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264101197-en. [7]

Poel, M. et al. (2015), *Data for Policy: A study of big data and other innovative data-driven approaches for evidence-informed policymaking*, Technopolis group, Oxford Internet Institute, Centre for European Policy Studies, http://www.data4policy.eu (accessed on 16 November 2018). [29]

Ribeiro, M., S. Singh and C. Guestrin (2016), ""Why Should I Trust You?": Explaining the Predictions of Any Classifier", http://arxiv.org/abs/1602.04938 (accessed on 29 November 2018). [59]

SAI India (2016), *Big Data Analytics in SAI India*, SAI India. [46]

Santiso, C. and B. Roseth (2017), "Data Disrupts Corruption", *Stanford Social Innovation Review*, Vol. 15/2. [21]

Schafer, B. (2016), "Surveillance for the masses: the political and legal landscape of the UK Investigatory Powers Bill", *Datenschutz und Datensicherheit - DuD*, Vol. 40/9, pp. 592-597, http://dx.doi.org/10.1007/s11623-016-0664-0. [6]

Tamkivi, S. (2014), *Lessons from the World's Most Tech-Savvy Government*, The Atlantic, https://www.theatlantic.com/international/archive/2014/01/lessons-from-the-worlds-most-tech-savvy-government/283341/ (accessed on 5 April 2017). [35]

The Centre for Open Data Enterprise (2018), *Roundtable on Leveraging Data as a Strategic Asset*. [69]

Tisné, M. (2018), *It's time for a Bill of Data Rights*, MIT Technology Review, https://www.technologyreview.com/s/612588/its-time-for-a-bill-of-data-rights/. [25]

Tuballa, M. and M. Abundo (2016), "A review of the development of Smart Grid technologies", *Renewable and Sustainable Energy Reviews*, Vol. 59, pp. 710-725, http://dx.doi.org/10.1016/j.rser.2016.01.011. [58]

Ubaldi, B. (2013), "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", *OECD Working Papers on Public Governance*, No. 22, OECD Publishing, Paris, https://dx.doi.org/10.1787/5k46bj4f03s7-en. [9]

Ubaldi, B. et al. (2019), "State of the art on the use of emerging technologies in the public sector", *OECD Working Papers on Public Governance*, No. 31, OECD, Paris. [15]

United Nations General Assembly Habitat III (2016), *Policy paper 3: National urban policy*, http://www.habitat3.org. [73]

United States White House (2018), *President's Management Agenda*, United States White House, https://www.whitehouse.gov/wp-content/uploads/2018/03/Presidents-Management-Agenda.pdf. [27]

Waag Society (2014), *Smart Citizen Kit Amsterdam - Meten is weten?*, https://waag.org/sites/waag/files/media/publicaties/eindrapportage-sck-asd.pdf. [49]

Welby, B. (2019), "The impact of Digital Government on citizen well-being", *OECD Working Papers on Public Governance*, No. 32, OECD, Paris. [14]

Weng, S. et al. (2017), "Can machine-learning improve cardiovascular risk prediction using routine clinical data?", *PLOS ONE*, Vol. 12/4, p. e0174944, http://dx.doi.org/10.1371/journal.pone.0174944. [1]