Oxford Commission on AI & Good Governance

OII OXFORD INTERNET INSTITUTE

UNIVERSITY OF OXFORD

Four Principles for Integrating AI & Good Governance

Written & Researched by Lisa-Maria Neudert & Philip N. Howard

Four Principles for Integrating AI & Good Governance

Written & Researched by Lisa-Maria Neudert & Philip N. Howard
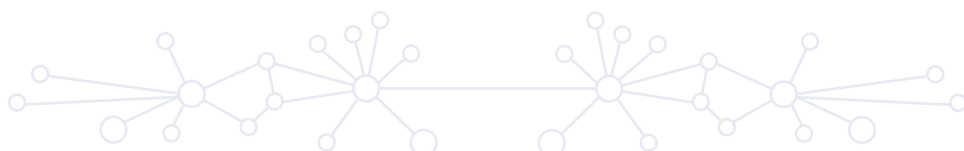
28 July 2020

Oxford Commission on
AI & Good Governance

Contents

Foreword

Artificial intelligence has already become deeply entrenched in public life. While the use of AI in governance is on the rise, pressing questions around its use in public service persist.

Increasingly, that means advanced machine learning tools are being adopted by governments and public agencies to tackle major public policy problems, rationalize the distribution of public services and goods, and automate the decision making and information flows of public institutions. How can we ensure that the opaque algorithms of machine learning systems used by public agencies are developed and deployed to protect human rights and democratic values? To what public problems can AI offer insights and solutions? Which domains of public policy should allow autonomous AI decision making, and under what circumstances? And what evidence do we need to assess the real-world impact of AI tools on public policy and good governance?

It is crucial that we plan now, for the implementation of AI systems we have currently, into the democratic governance systems we want to strengthen. In a wide range of applications AI systems have been found to manifest existing biases, amplify racial inequalities, and discriminate against the underprivileged. Research has demonstrated how AI systems can simplistically reduce real-world complexities, and can be leveraged for political manipulation, individual surveillance, and social control.

The ongoing COVID-19 pandemic has supercharged the pressure for widespread surveillance, data collection, and the use of AI to study and solve social problems. As nations around the globe rush to deploy contact tracing applications, it has become abundantly clear that governments require new regulatory frameworks and expertise to develop new technologies that adhere to democratic values and human rights.

Now more than ever, governments need to adopt AI as a force for change and justice. But unfortunately, many of the current systems manifest and exacerbate policy problems. And they do so widely uncontested, because the guidelines for integrating AI and good governance systems have yet to be set.

Weighing opportunities on the one hand against profound risks on the other, governments struggle with complex questions about the impact of AI-driven tools in public service. Trapped in indecision and uncertainty, governments risk missing opportunities and falling behind technological progress. Integrating AI tools into public administration, however, must be done with care. As governments turn to AI to solve policy problems, we need rules in place to safeguard democratic values.

That is why we are launching the Oxford Commission on AI & Good Governance (OxCAIGG): to unite experts on governance, technology, security, and human rights to envision comprehensive guidelines and policy research on the use of AI in public administration. Over the next eighteen months OxCAIGG will produce actionable, evidence-based recommendations to enable governments and public administrations to harness the opportunities of AI while protecting democracy.

Philip N. Howard

Chair & Commissioner, Oxford Commission on
AI & Good Governance
Director of the Oxford Internet Institute

Executive Summary

Artificial intelligence (AI) has arrived in the public sector. Promising to solve some of our most pressing public policy and societal challenges, AI will become ever more pervasive.

Many governments, public agencies and institutions already employ AI in providing public services, the distribution of resources and the delivery of governance goods. In the public sector, AI-enabled governance may afford new efficiencies that have the potential to transform a wide array of public service tasks.

But short-sighted design and use of AI can create new problems, entrench existing inequalities, and calcify and ultimately undermine government organizations. Frameworks for the procurement and implementation of AI in public service have widely remained undeveloped. Frequently, existing regulations and national laws are no longer fit for purpose to ensure good behaviour (of either AI or private suppliers) and are ill-equipped to provide guidance on the democratic use of AI.

As technology evolves rapidly, we need rules to guide the use of AI in ways that safeguard democratic values. Under what conditions can AI be put into service for good governance?

We offer a framework for integrating AI with good governance. We believe that with dedicated attention and evidence-based policy research, it should be possible to overcome the combined technical and organizational challenges of successfully integrating AI with good governance. Doing so requires working towards:

I **Inclusive Design**: issues around discrimination and bias of AI in relation to inadequate data sets, exclusion of minorities and under-represented groups, and the lack of diversity in design.

II **Informed Procurement**: issues around the acquisition and development in relation to due diligence, design and usability specifications and the assessment of risks and benefits.

III **Purposeful Implementation**: issues around the use of AI in relation to interoperability, training needs for public servants, and integration with decision-making processes.

IV **Persistent Accountability**: issues around the accountability and transparency of AI in relation to 'black box' algorithms, the interpretability and explainability of systems, monitoring and auditing.

Over the next eighteen months OxCAIGG will develop actionable, evidence-based recommendations to enable governments and public administrations to harness AI. Our mission is to:

1 **Act** with agility and purpose to understand the impact of the AI tools currently being used in governance, public administration, securing social welfare and providing public goods.

2 **Assist** policymakers and entrepreneurs with policy questions, public applications, and design ideas that put AI and machine learning into public service.

3 **Identify** solutions and guide policymaking processes in a way that strengthens the development pathway for bringing AI to solve rather than complicate social problems.

1 Introduction

Artificial intelligence (AI) is transforming our world, and this transformation is happening now.

Artificial Intelligence is already shaping public and private life. It is not a technology of the future, an imagined application of new tools, or a hypothetical means of solving some of grand challenges we face. AI, machine learning and other advanced algorithmic systems are helping to manage our homes and workplaces and creating both opportunities and constraints on our lives. The organisations driving the widespread adoption of AI technologies increasingly include governments, public agencies and the wider public sector who use AI for the provision of public services, the distribution of resources and the delivery of governance goods.

AI promises to solve some of our most pressing and persistent public policy challenges. From economic growth to public welfare, to product design to institutional innovation, and human rights to social mobility, machine learning applications have enormous potential to redesign and reconfigure how we get things done. When put to use in the public sector, AI-enabled policymaking and governance may afford new efficiencies that have the potential to transform a wide array of public service tasks.[1]

However, with these opportunities come profound challenges. Short-sighted design, procurement and implementation of AI can create new problems that entrench existing inequalities, perpetuate discriminatory practices, calcify and ultimately undermine government organizations.[2] What is more, AI-driven decision-making is often locked away in algorithmic "black boxes" that escape explanation or understanding, and obscure the rationales leading to a given outcome.[3]

Governments are moving quickly to put AI to the task of public service. But at the same time, they are also struggling to adopt and adapt these technologies as existing frameworks for procurement and implementation are no longer fit for purpose and ill-equipped

to provide guidance on the democratic use of AI. While there is growing awareness of the ethical challenges our societies face as AI develops, guidelines for the procurement and use of AI in public service have remained underdeveloped. When it comes to the governance of new technologies, the views of public and private actors often diverge across different geopolitical contexts. [4] As technology evolves rapidly and the potential of AI for good governance is great, we need rules to guide the development, procurement and use of AI in ways that safeguard democratic values.

In this working paper, we develop a comprehensive framework of governance issues in relation to the use of AI in public service as a starting point for policy action. We begin by offering a policy portfolio of the plethora of present and potential uses of AI for good governance. We then lay out the four most pressing policy challenges around AI that governments must address now to ensure the democratic use of AI for good governance: (I) inclusive design; (II) informed procurement; (III) purposeful implementation; (IV) persistent accountability.

The thinkpiece we present here, marks the first output of the Oxford Commission on AI & Good Governance, OxCAIGG, and offers an agenda for our priorities over the next eighteen months. In an effort to address the four challenges outlined here, OxCAIGG will bring forward concrete policy recommendations targeting the government development, procurement, and implementation of AI-driven technologies. Our recommendations will build on original research and empirical evidence, briefings with stakeholders from governments, public agencies, industry and civil society and the expansive expertise and experience of our OxCAIGG Commissioners.

2 Good Governance with an AI Toolkit

Decisions about who to lend money to, who to provide insurance for, how to price consumer goods, even who to employ: all of these decisions can now be supported (if not entirely made) by artificial intelligence and decision support systems.

Artificial intelligence has revolutionised the way the private sector does business in data-rich markets.[5] Autonomous agents such as chatbots are starting to come into the front lines of customer service. These technologies are now also beginning to make their way into government: decisions about parole, about child welfare, about adult social care, and social benefits—all of these are potential areas where artificial intelligence can contribute.[6]

The emergence of these technologies poses huge questions for what modern governance is and should be. Historically, new information technologies have brought about tectonic shifts in the administration of collective welfare and public goods. Over the last century, bureaucratic authority has moved from the 'street level' of local government and neighbourhood administrative services to national and 'system level' services from more centralized government offices.[7]

This new wave promises a shift to the 'algorithm level'. Such a change could mean further loss of administrative autonomy, not just by frontline public servants but also by managers, who may struggle to understand the decisions being produced by difficult-to-explain computational techniques. However it could also mean that such decision makers also gain more power, as they are pushed into new situations of evaluating and controlling automated decision systems.[8]
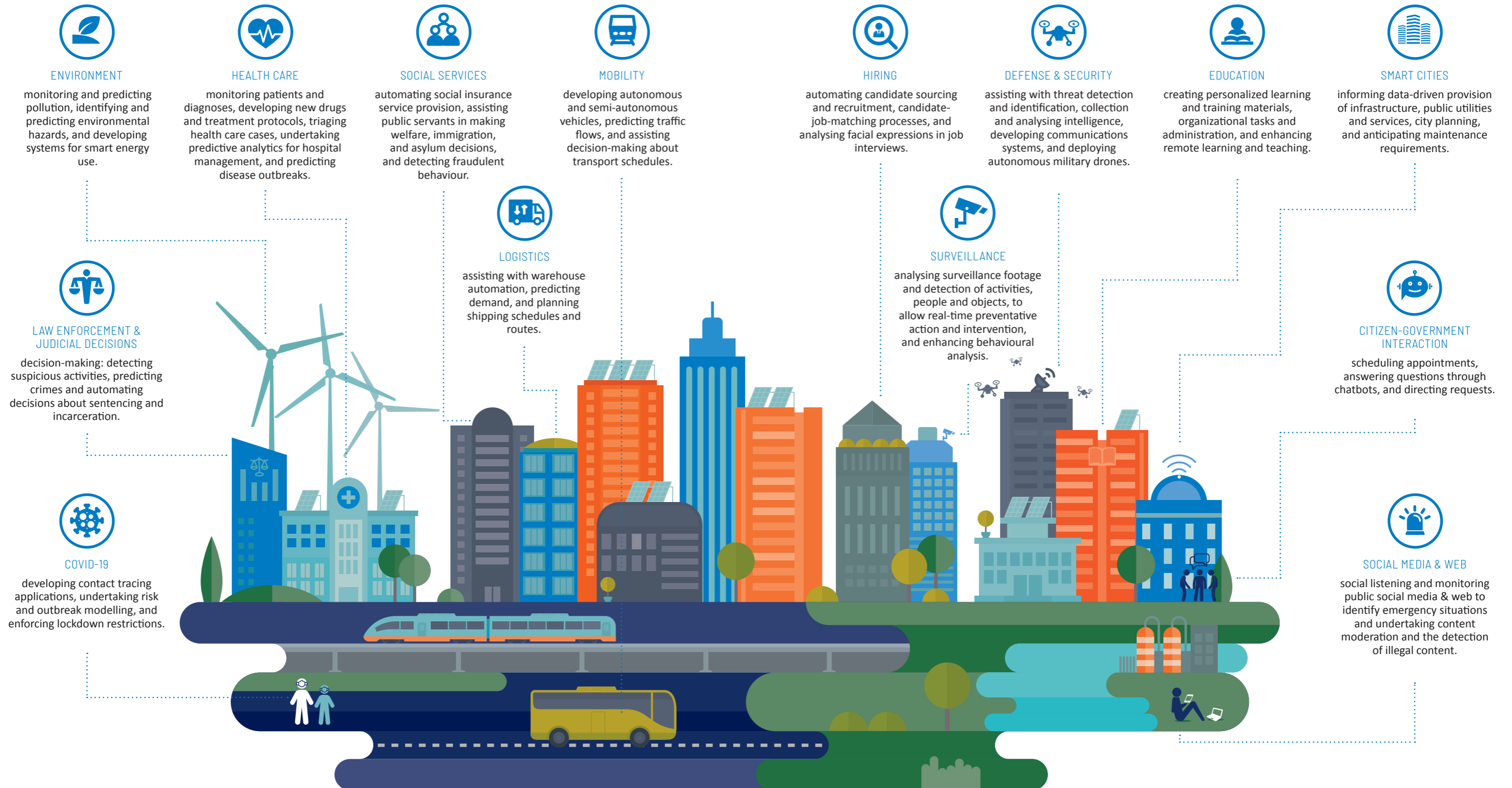
There are also important consequences for citizens. First, algorithms can generate arbitrary and unjustified decisions about individuals, who are then left with insufficient legal remedy to contest these decisions from systems they struggle to understand.[9],[10] Second, they widen the scope of potential groups subjected to discriminatory effects who lack appropriate protection in existing anti-discrimination law.[11] Third, they make the decision-making process more opaque, by preventing the allocation of responsibility under existing accountability legislation or guidelines.[12]

In addition to all of this, many government departments and agencies - especially at the local level - remain uncertain about how to engage with artificial intelligence and other automated decision support technologies. Lacking guidelines and best practices, governments are at risk of being sold unsuitable technologies by consulting and IT firms or developing inadequate products in-house.

There are many domains of governance in which AI now have a role. We offer a comprehensive portfolio of policy-relevant present and potential uses of AI for governance.

POLICY-RELEVANT PRESENT AND POTENTIAL USES OF AI FOR GOVERNANCE

ENVIRONMENT

monitoring and predicting pollution, identifying and predicting environmental hazards, and developing systems for smart energy use.

HEALTH CARE

monitoring patients and diagnoses, developing new drugs and treatment protocols, triaging health care cases, undertaking predictive analytics for hospital management, and predicting disease outbreaks.

SOCIAL SERVICES

automating social insurance service provision, assisting public servants in making welfare, immigration, and asylum decisions, and detecting fraudulent behaviour.

MOBILITY

developing autonomous and semi-autonomous vehicles, predicting traffic flows, and assisting decision-making about transport schedules.

HIRING

automating candidate sourcing and recruitment, candidate-job-matching processes, and analysing facial expressions in job interviews.

DEFENSE & SECURITY

assisting with threat detection and identification, collection and analysing intelligence, developing communications systems, and deploying autonomous military drones.

EDUCATION

creating personalized learning and training materials, organizational tasks and administration, and enhancing remote learning and teaching.

SMART CITIES

informing data-driven provision of infrastructure, public utilities and services, city planning, and anticipating maintenance requirements.

LOGISTICS

assisting with warehouse automation, predicting demand, and planning shipping schedules and routes.

SURVEILLANCE

analysing surveillance footage and detection of activities, people and objects, to allow real-time preventative action and intervention, and enhancing behavioural analysis.

LAW ENFORCEMENT & JUDICIAL DECISIONS

decision-making: detecting suspicious activities, predicting crimes and automating decisions about sentencing and incarceration.

CITIZEN-GOVERNMENT INTERACTION

scheduling appointments, answering questions through chatbots, and directing requests.

COVID-19

developing contact tracing applications, undertaking risk and outbreak modelling, and enforcing lockdown restrictions.

SOCIAL MEDIA & WEB

social listening and monitoring public social media & web to identify emergency situations and undertaking content moderation and the detection of illegal content.

3 Four Principles of AI for Good Governance

Certainly, an information-rich governance system could use AI and machine learning tools to support and extend a deliberative democracy. The information available and interpretable by policy makers is often incomplete, imprecise, false, too complex, or too reductive.

Often the most intractable problems of deliberative democracy are framed as problems of information quality or quantity.[13] Unfortunately, as the computational toolkit for redressing these shortcomings has improved, the social institutions to support these imagined possibilities are rarely formed at the same pace.[14] In other words, the arrival of AI systems may become an occasion for organizational restructuring in public administration itself.

One lesson from the history of technology in government is that we must act purposefully in designing technology from the bottom up to serve democratic institutions. Despite all the principled and abstract discussions about how AI can be designed to improve public life, it will take the right regulatory, economic, and experiential context to integrate AI fully into public service.

Good governance can mean different things in different contexts, but at its simplest it refers to the ability of public agencies to provide for the public good in efficient, effective, and sustainable ways. If AI systems can be purposefully set to work for good governance, especially in democracies, it will be by overcoming four significant challenges in its development and application.

I INCLUSIVE DESIGN

The public must have sight of how public-service AI is designed, and AI systems must be built with everyone in mind.[15] AI systems have been shown to reproduce imbalances and discriminate against minorities and under-represented groups in the population. Typically, an AI forms such biases when it is processing data sets that are insufficiently or inaccurately representing certain groups and conditions, or when data sets reflect existing injustices and other grievances. For example, a machine learning algorithm calculating credit limits may discriminate against women because historically credit systems have been biased towards favouring men.

The public must have sight of how public-service AI is designed, and AI systems must be built with everyone in mind.

Principles of equal treatment are built into the constitutions and charters of many democracies, and AI systems need to abide these rules. Best practices and guidelines for addressing bias in AI are already widely available, though these rarely sufficiently adhered

to in practice.[16] Governments and public agencies need to review flaws and biases within data sets carefully and where necessary address insufficiencies or inaccuracies, for example by building in parameters that help AI systems to interpret the information it has been given.

The scale of this task is enormous. AI-enabled public services will make use of a wealth of publicly held and privately held databases about personal behaviours, attitudes, and perceptions and ultimately, that data will come from a diverse range of platforms, devices and public networks. The use of personal data needs to be lawful, ethical, and secure, and, what is more, it must be inclusive. Inclusivity also extends to the development of AI itself.

Currently, the field of AI for democratic governance is dominated by a handful of powerful US-centric technology companies, that frequently skew towards dominantly white and male leadership teams. Diverse and multidisciplinary teams mitigate risks of bias and reflect real-world conditions. At a time when discrimination, racial inequalities, and hyper-partisan bigotry are surging, AI must be leveraged to promote equality. AI needs to advance inclusivity, not perpetuate systemic injustices.

How can policymaking guide good practices for the use of private and public data for inclusive AI? What kinds of policy interventions might improve the ways technology can help overcome some of society's most intractable problems of inequality? How can we improve the diversity of voices in the design of AI tools?

II INFORMED PROCUREMENT

Public agencies need to be trained up for the complex questions that arise in AI procurement, and AI procurement should not proceed without a full understanding of data provenance, modelled outputs, and the governance structure of the firms producing the AI. To harness artificial intelligence in the public service, governments need to take decisions about their procurement and development. When obtaining AI-driven solutions from external suppliers or developing custom-made AI-tools, governments need to consider questions around the design, usability, and aptitude of AI alongside other considerations around budget and timelines.

> *Public agencies need to be trained up for the complex questions that arise in AI procurement, and AI procurement should not proceed without full understanding of data provenance, modelled outputs, and the governance structure of the firms producing the AI.*

Governments need to assess the risks and benefits of AI applications, whether they offer a solution to a set of problems, and whether such technologies operate legally, ethically, and inclusively.[17] Assessing the risks and benefits of deploying AI-technologies is also necessary. However, the procurement and development of AI poses several new challenges to governments. Complex technologies and their impacts are extremely difficult to evaluate and to conduct due diligence for.

3. Four Principles of AI for Good Governance continued

Frequently, algorithms are proprietary and thus remain opaque. What is more, assessing cutting-edge technology requires sophisticated expertise and specific knowledge which is not available to many governments.

In addition, in order to deploy AI-tools public administrations need to engage effectively with a wide array of AI vendors, including big tech, start-ups, and research institutions, or build these capacities in-house. Already, several governments have guidelines for procuring or developing technology, the use of sensitive data, data transportability and security, but best practices and regulatory frameworks frequently fall short of addressing issues specific to AI and its impact on policy decisions.

How can we equip public administration with the tools and skills to evaluate that AI-driven technologies are lawful, ethical, and functional? What guidelines, best practices and regulatory frameworks are needed to enforce good behaviour? How can government procurement systems support industry to create systems designed for good governance?

> *Putting AI-driven applications into use in public service requires public servants to ask and answer many kinds of questions about their implementation and integration with existing practices and systems.*

III PURPOSEFUL IMPLEMENTATION

Putting AI-driven applications into use in public service requires public servants to ask and answer many kinds of questions about their implementation and integration with existing practices and systems. Once acquired, governments must follow through with the necessary staff allocations, training and adaptations in decision-making processes. As machine learning and automated systems become sophisticated enough to operate with some level of autonomy, governments need to understand how decisions made by AI tools shape policy outcomes. What is more, governments need to assess in which policy areas AI systems are capable of good governance and can be implemented to begin with.

Already, there are many uses of AI in the public service, where AI is used to automate tasks, for example for vetting CVs of job applicants or classifying emergency calls based on their urgency. Frequently, there is still a human-in-the-loop to train, test and supervise AI systems and intervene when a system returns unreliable or problematic results. However, there is little guidance available on automated versus human decision-making, for example on under what conditions AI systems may act autonomously or how disagreements between AI and human decision-making are reconciled. What is more, there is little evidence available to analyse the impact of AI-based decision making as compared to human decision-making or human-supervised decision-making.

AI can automate and make more efficient otherwise laborious, repetitive, and even complex tasks, but we need rules to ensure that AI is implemented and integrated in ways that advance the

public good. There are also several more practical questions in relation to the use of AI, for example around the interoperability of new tools with existing systems, the transferability of data or technical requirements for hardware and software. Adopting new systems frequently requires technological and organizational adaptations including training which should be considered when deploying AI.

Under what conditions can AI autonomously automate policy relevant decision making and when is a human-in-the-loop required? How do humans and AI systems co-operate to make decisions? To what policy problems does AI offer solutions? What guidelines are required to ensure that AI tools can be integrated with existing practices and systems?

IV   PERSISTENT ACCOUNTABILITY

Once designed, procured, and implemented, the new organizational and technical systems must be regularly evaluated and appraised, both for unexpected consequences and new opportunities. AI systems are by design complex, frequently operate on opaque algorithms and rely of vast amounts of data to learn and operate. As a result, AI tools are often labelled as 'black boxes' that return obscure outputs that largely evade accountability. Democratic policy making needs to fulfil standards of accountability towards the public, and as such AI systems also need to comply with such standards. To establish accountability, AI systems, data, processes, and outputs need to be made transparent. This also extends to proprietary algorithms when governments and public agencies use AI tools from external suppliers.

Once designed, procured and implemented, the new organizational and technical systems must be regularly evaluated and appraised, both for unexpected consequences and new opportunities.

To ensure transparency, governments should require explainability and interpretability of AI algorithms and their design.[18], [19] Tiered models are conceivable: For example, anyone operating an AI-enabled tool should be able to grasp basic functionalities–even without any prior technology expertise–whereas a deeper understanding is required of technology experts in government. At all times, AI systems should make transparent how and why a model performed in a specific way. What is more, governments and external supplier should be obliged to make transparent system updates, major bug fixes and security breaches including hacks and leaks.

Furthermore, AI technologies need to adhere to public expectations for accountability. Implementing oversight and review processes, such as external audits or process logs of activity, and requiring of suppliers the openness and transparency to make their systems auditable, are potential tools to enforce good behaviour. Governments could also consider encouraging open source practices and peer review processes to increase transparency.

What transparency and accountability systems should be required of AI systems, their processes, and outcomes in public service? What tools and standards are needed to accomplish accountability of AI tools in governance? How can we measure the social impact of AI and what would such audits look like?

4 Conclusion

Artificial intelligence is arriving now and has already become integrated in many areas of public service. Offering new efficiencies and transformative applications, AI will expand and proliferate as the technology advances and becomes continuously more accessible.

But in the face of crisis and emergency, the need for AI for governance may suddenly become urgent–and governments are ill-equipped to make good decisions on behalf of the public. And as AI is quick to advance, we need frameworks that are future-proof and robust to technological innovation.

The COVID-19 pandemic has revealed deficiencies in government capacities around the use of new technologies, including artificial intelligence.[20] While technology experts have pointed to numerous opportunities to adopt AI-enabled tools for combatting COVID-19, including contact tracing and outbreak modelling, governments have struggled to put AI into use.[21] Even the most advanced democracies have grappled with developing frameworks for the use of AI that conform with national laws and public expectations. Finally, developing these tools in ways that conform with democratic values has proven challenging – but governments less concerned with issues around privacy, surveillance and data protection have leveraged AI against COVID-19 with some success.

Faced with the sheer potential of AI-driven public service technologies, governments are pressed for time for devising frameworks to put AI into use to meet an urgent need. Even now governments are missing out on opportunities to harness AI for solving pressing policy problems that are otherwise insoluble or require tremendous resources. And while democratic governments are stuck developing policy processes that are notoriously time-consuming, authoritarian regimes are already capitalising on AI for political and economic gain and control.

What tools and measures are available to governments to effectively ensure good governance when facing urgent AI implementation needs? How can governments effectively incentivize and encourage the development of AI-enabled applications for governance? What kinds of policy interventions are future-proof against rapid technological innovation?

Artificial intelligence is arriving now and has already become integrated in many areas of public service. Offering new efficiencies and transformative applications, AI will expand and proliferate as the technology advances and becomes continuously more accessible.

5 Objectives for the Oxford Commission on AI and Good Governance

While there is excitement about the prospects for AI in public service, there are strong public concerns and fears for how it will be used in governance. Now, we need evidence-based recommendations and research to advance the conversation.

We almost certainly want the potential benefits of logical decision making, problem solving, resource allocation and economic efficiency in governance, without the risks of perpetuating social inequalities, losing political accountability, and causing additional public policy problems. We need to investigate the consequences of deploying AI-driven decision support systems in local and national government.

AI holds great promise for solving public problems and strengthening deliberative democracy and good governance. However, this is just potential, and the purposeful application of AI in governance is not simply about philosophical puzzles or hypothetical scenarios—it requires evidence-based policy analysis about the impact of AI on citizens and civic engagement, and insights into the transportable best practises that will improve deliberative democracy at all government levels.

The application of AI tools to public policy problems has already arrived and the use of AI will continue to surge as governments leverage readily available data for efficiency. Governments and public agencies are among the first organisations to employ AI tools for public welfare, and we already have implementation cases from which to learn and develop best practices.

When sophisticated AI-driven algorithms have a role in determining the provision of government health services, evaluating our tax burden, or assessing security risks they are providing basic governance goods. But under what conditions would we want

AI providing governance goods and how can we ensure good governance? OxCAIGG's aim is to provide evidence guidance on pressing policy issues around AI in public service. Our mission is to:

1. **Advance** policy research into the credible, trustworthy role for AI and machine learning tools in how governments make and implement public policy

2. **Provide** a "rapid response" capacity to assist policy makers and public officials in times of political sensitivity or unique opportunity in using AI to solve collective problems

3. **Identify** solutions, best practises, and policy options that guide policy making processes in a way that strengthens governance systems and bolsters democratic values.

Now is the time to ask and answer questions about the conditions under which governments and public agencies can practically use AI for providing public services both now and in a post COVID-19 world. Public agencies evaluate new AI applications for providing governance goods, but do not have the policies in place for evaluating, procuring, and implementing AI tools for public service. The goal of the Oxford Commission on AI and Good Governance is to develop policies urgently needed to ensure the democratic use of AI for good governance.

Acknowledgements

References

[1] Margetts, H. & Dorobantu, C. Rethink Government with AI. *Nature* 163–165 (2019) doi:10.1038/d41586-019-01099-5.

[2] Noble, S. *Algorithms of Oppression: How Search Engines Reinforce Racism*. (NYU Press, 2018).

[3] Howard, P. N. *Pax Technica: The Impact of Automation on Public Opinion*. (Yale University Press, 2015).

[4] O'Hara, K. & Hall, W. Four Internets. *Commun*. ACM 63, 28–30 (2020).

[5] Mayer-Schönberger, V. & Ramge, T. *Reinventing Capitalism in the Age of Big Data*. (Basic Books, 2018).

[6] Leslie, D., Holmes, L., Hitrova, C. & Ott, E. *Ethics of Machine Learning in Children's Social Care*. (2020).

[7] Bovens, M. & Zouridis, S. From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. *Public Adm. Rev*. 62, 174-184 (2002).

[8] Vogl, T., Seidelin, C., Ganesh, B. & Bright, J. Algorithmic Bureaucracy. in *Proceedings of the 20th Annual International Conference on Digital Government Research* 148–153 (Association for Computing Machinery, 2019). doi:10.1145/3325112.3325240.

[9] Wachter, S., Mittlestadt, B. & Floridi, L. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. Int. *Data Priv. Law* 7, 76–99 (2017).

[10] Veale, M., Van Kleek, M. & Binns, R. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1–14 (Association for Computing Machinery, 2018). doi:10.1145/3173574.3174014.

[11] Barocas, S. & Andrew, D. Big data's disparate impact. *Calif. Law Rev*. 104, (2016).

[12] Pasquale, F. A. *The Black Box Society: The Secret Algorithms That Control Money and Information*. (Harvard University Press, 2015).

[13] Dahl, R. A. *Democracy and its critics*. (Yale University Press, 1989).

[14] Barley, S. Technology As An Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments. *Adm. Sci*. Q. 31, 78–108 (1986).

[15] Daugherty, P., Wilson, J. & Chowdhury, R. Using Artificial Intelligence to Promote Diversity. *MIT Sloan Manag. Rev*. 60, (2019).

[16] OECD. *Recommendation of the Council on Artificial Intelligence*. (2019).

[17] Leslie, D. Understanding Artifical Intelligence Ethics and Safety: A Guide for the Responsible Design and Implementation of AI Systems in the Public Sector. (2019) doi:https://doi.org/10.5281/zenodo.3240529.

[18] Mittelstadt, B., Russell, C. & Wachter, S. Explaining Explanations in AI. in *Proceedings of the Conference on Fairness, Accountability, and Transparency* 279–288 (2019). doi:https://doi.org/10.1145/3287560.3287574.

[19] ICO & The Alan Turing Institute. *Explaining Decisions Made with AI*. 136 (2020).

[20] Howard, P. N. & Neudert, L.-M. AI can battle coronavirus, but privacy shouldn't be a casualty. *TechCrunch* (2020).

[21] Leslie, D. Tackling COVID-19 through Responsible AI Innovation: Five Steps in the Right Direction. *Harv. Data Sci. Rev*. (2020).

About the Authors

Lisa-Maria Neudert is a DPhil candidate at the Oxford Internet Institute and a researcher with the Computational Propaganda Project, where her work is located at the nexus of political communication, technology studies, and governance. Her current research is looking into the public and private governance of policy issues surrounding disinformation through governments and social media platforms. She is the Co-Principal Investigator for the project's research on Europe supported by Civitates and the Secretary to the Oxford Commission on AI & Good Governance, OxCAIGG. Neudert holds an MSc in Social Science of the Internet from the University of Oxford, and a BA in Communication Science from the Ludwig-Maximilians-University in Munich. Selected as a Fulbright scholar, she previously studied at Georgetown University and the National University of Singapore. Neudert's commentary writing has appeared in several outlets including The Guardian, the MIT Tech Review and TechCrunch. Her research has been featured in several media outlets including the Washington Post, the Financial Times and Der Spiegel. She has presented her research at the European Commission, The Council of Europe and The World Economic Forum in Davos. Neudert is on Twitter @lmneudert.

Philip N. Howard is a professor and writer. He teaches at the University of Oxford and directs the Oxford Internet Institute. He writes about information politics and international affairs, and he is the author of ten books, including *The Managed Citizen, Pax Technica*, and *Computational Propaganda.* He has won multiple best book awards, and his research and commentary writing has been featured in the New York Times, Washington Post, and many international media outlets. Foreign Policy magazine named him a "Global Thinker" for 2018 and the National Democratic Institute awarded him their "Democracy Prize" for pioneering the social science of fake news. He has testified before the US Senate, UK House of Parliament, and European Commission on the causes and consequences of fake news and misinformation. His latest book is *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. He blogs at www.philhoward.org and tweets from @pnhoward.

About the Oxford Commission on AI and Good Governance

The mission of the Oxford Commission on AI and Good Governance (OxCAIGG) is to investigate the artificial intelligence implementation challenges faced by governments around the world, identify best practices for evaluating and managing risks and benefits, and recommend strategies for taking full advantage of technical capacities while mitigating potential harms of AI-enabled public policy. Drawing from input from experts across a wide range of geographic regions and areas of expertise, including stakeholders from government, industry, technical and civil society, OxCAIGG will bring forward applicable and relevant recommendations for the use of AI for good governance.

oxcaigg.oii.ox.ac.uk