Good Practice Principles for Data Ethics in the Public Sector

ACKNOWLEDGEMENTS

This document was prepared by the OECD Digital Government and Data Unit within the Open and Innovative Government Division (OIG) - Public Governance Directorate (GOV). Jacob Arturo Rivera Pérez, Policy Analyst, Digital Government and Data, OECD, co-drafted this document and served as the lead coordinator for this project. Lucia Chauvet, Jr. Policy Analyst, Digital Government and Data, OECD, provided data analytical work and project management support. Barbara-Chiara Ubaldi, Head of the Digital Government and Data Unit (GOV/OIG), provided strategic orientation and revisions to the document.

The team is especially grateful to Jaron Haas, Simone Schoof and Marieke Schenk from the Netherland's Ministry of the Interior and Kingdom Relations (MINBZK) for their leadership and the support provided to the OECD during the preparation of these Good Practice Principles. The OECD Secretariat acknowledges the invaluable contribution of Natalia Domagala (Government Digital Service, United Kingdom) and Omar Bitar (Treasury Board of Canada Secretariat, Canada). The OECD thanks all public officials from Australia, Argentina, Belgium, Brazil, Chile, Colombia, Denmark, Finland, Greece, Israel, Italy, Ireland, Latvia, Lithuania, Luxembourg, Poland, Portugal, Panama, Slovenia, South Korea, Spain, Sweden and Uruguay for their participation in the activities of the Thematic Group on Data-driven Public Sector from October 2018 to December 2020, and for their valuable comments to the document.

The team also thanks the following colleagues and partners for their comments and revisions to the document: Elettra Ronchi, Christian Reimsbach-Kounatze, Karine Perset and Luis Aranda (Science, Technology and Innovation Directorate, OECD); Rasmus Raabjerg Nielsen and Anna Byhovskaya (Trade Union Advisory Committee, TUAC, to the OECD); Julio Bacio Terracino (Public Sector Integrity Division, GOV); Stéphane Jacobzone and Lizeth Fuquene (Governance Reviews and Partnership Division, GOV); Costanza Caputi (Infrastructure and Public Procurement Division, GOV); Gregor Virant and Nick Thijs (SIGMA, GOV); Miguel Amaral and Anna Pietikainen (Regulatory Policy Division, GOV); Jamie Berryhill and Kent Aitken (Observatory of Public Sector Innovation, OIG, GOV); Emma Cantera and Marie Whelan (Open Government and CivicSpace Unit, OIG, GOV); Lawrence Pacewicz (OECD Legal Department); Thom Townsend (Open Ownership); Lindsey Marchessault (Open Contracting); and Juan Pablo Guerrero and Lorena Rivero del Paso (Global Initiative for Fiscal Transparency, GIFT).

For more information please contact: Barbara.UBALDI@oecd.org; JacobArturo.RIVERAPEREZ@oecd.org; Lucia.CHAUVET@oecd.org

Introduction

By contributing to the development of commonly shared good practice principles for data ethics[1] in the public sector, OECD Member and non-Member countries can advance towards a multilateral human-centred digital government and data agenda drawing upon their role as digital leaders. This, by taking common actions that place human rights and values at the core of digital government and data policies, strategies, projects and initiatives.

Increasing data flows across governments, sectors and borders demand commonly defined data governance frameworks and the promotion of greater multilateral and multi-stakeholder digital co-operation and collaboration. Yet, given the plethora of existing data regimes[2] in the world, the complexity of defining and agreeing upon meta-governance structures for data (e.g. cross-border data governance arrangements) highlights the need to provide more granular policy guidance on the ethical implications of accessing, sharing and using data.

Data environments are increasingly complex. The movement or transfer of data across environments presents new challenges (e.g. data integration or analytics may erode privacy protections; data users may unknowingly violate community controls). Therefore, agreeing on a basic and common guidance on data ethics in the public sector can help ensure that relevant rules and procedures "move with the data" and prevent inadvertent breaches of ethical principles in practice.

Societal demand for ethical practices to complement data protection and privacy regulations has increased, reflecting a growing interest in ensuring that data is used in ways that respect the public interest and deliver trustworthy outcomes.

The emergence of new technologies[3] and their increasing uptake in governments has coincided with an exponential increase in data generation and use resulting from digitalisation, which has in turn expanded the possibilities of data analytics. Governments are increasingly leveraging digital technologies to improve and streamline core government functions, inform the design and delivery of better policies and services and, where feasible and appropriate, automate decision-making using algorithms to process this data at scale.

While some laws and regulations set rules for the protection, management and publication of data, the development of values-based guidance such as data ethics frameworks provides governments with an opportunity to:

- Adopt inclusive and collaborative approaches to designing data policies, strategies and initiatives that reinforce the ethical use of data in the public sector;
- Build consensus on how to foster public trust in practice in the context of data access, sharing and use; and
- Agree on trustworthy data management practices that adhere to shared values, at both operational and strategic levels.

The ethical use of data in the public sector calls for embedding ex-ante and ex-post risk-management approaches in order to address hazards and trade-offs. In practice, data ethics should translate into specific actions throughout the data value cycle[4].

The nature and diversity of data typologies, taxonomies and formats (e.g. research data, administrative data, national statistics, health data, non-personal vs. personal data, aggregated vs. granular data, structured vs. unstructured data) add to the complexity of the policies and data governance arrangements needed to enhance their trustworthy management across the different stages of the data value cycle. These stages include, but are not restricted to, data generation, selection, collection, curation, storage, disposal, access, sharing, and use.

For example, governments need to be prepared to take action to address issues and concerns associated with data corruption; biases affecting the generation of data or its extraction (e.g. selection of data sources); and the quality of data inputs used to train Artificial Intelligence (AI) models. Other hazards include data misuse and abuse by individuals and organisations and the delivery of negative outcomes through data use, including in the context of AI systems[5].

Objectives

The *Good Practice Principles for Data Ethics in the Public Sector* (hereafter, 'Good Practice Principles') presented in this paper seek to shed light on the value and practical implications of data ethics in the public sector. They aim to support public officials in the implementation of data ethics in digital government projects, products, and services such that i) trust is placed at the core of their design and delivery and ii) public integrity is upheld through specific actions taken by governments, public organisations and, at a more granular level, public officials.

These Good Practice Principles emerge from observed practices in digital government and data-driven public sectors across OECD Member and non-Member countries. They intend to support public officials in countries that have adhered to the OECD Recommendation of the Council on Digital Government Strategies [OECD/LEGAL/0406] in implementing its provisions, namely provision 3.

At the same time, these Good Practice Principles reinforce the policy measures and values-based approaches addressed in i) a number of existing OECD legal instruments, including the Recommendation on Public Integrity [OECD/LEGAL/0435], the Recommendation on OECD Guidelines for Managing Conflict of Interest in the Public Service [OECD/LEGAL/0316], the Recommendation on Artificial Intelligence [OECD/LEGAL/0449]), the 2013 OECD Privacy Guidelines[6]; and ii) the OECD work done in the context of the development of the draft Recommendation on Enhancing Access to and Sharing of Data[7].

The Thematic Group on Data-driven Public Sector, meeting under the aegis of the OECD Working Party of Senior Digital Government Officials (E-leaders), has drawn together these Good Practice Principles for Data Ethics in the Public Sector. For details on the methodology for their development see Annex A.

Considerations relevant to the development of the Good Practice Principles

OVERARCHING CONSIDERATIONS

In developing the Good Practice Principles for Data Ethics in the Public Sector, the Thematic Group on Data-driven Public Sector acknowledged that:

- Data ethical frameworks do not replace, but rather complement, support, and are interconnected with relevant hard law instruments such as regulations on privacy, data protection, open data, open government, transparency and data sharing within the public sector, among others.
- The publication of and adherence to non-binding guidelines or standards such as the Good Practice Principles do not guarantee real-world implementation. The effective alignment with and success of data ethical frameworks require their incorporation into public sector decision-making processes and the articulation of specific actions at a more granular and technical level (e.g. data management rules).
- The effectiveness of data ethical frameworks is not achieved in isolation. Putting in place sound data governance arrangements[8] in the public sector (e.g. institutional roles and responsibilities, co-ordination fora, advisory bodies, and accountability mechanisms) is a precondition for success.
- Data ethics in the public sector should connect organically with broader policies and the tools and mechanisms derived from them. These include policies on public integrity (e.g. ethics, conflict of interest, auditing, whistleblower protection), digital government (e.g. decisions on ICT project funding, procurement of digital projects, service standards, data governance and management, citizen-centricity, open data, open source code), open government (e.g. proactive transparency of government decisions, access to information, stakeholder engagement, deliberative democracy, public communication), public sector employment (diversity in the workplace), and social inclusion (including digital inclusion and digital rights).
- Data governance and AI governance can intersect at their strategic, tactical and technical layers. This intersection creates clear synergies between data ethics and AI ethics and underscores the values-based approaches shared across these areas - both in terms of their conception and application. However, while these Good Practice Principles highlight those synergies where relevant, their main objective is to underline issues that are specific to data ethics, including in the context of AI systems and automated decision-making in the public sector.

In developing the Good Practice Principles, the Thematic Group on Data-driven Public Sector reiterated the particular relevance of the following cross-cutting issues:

Data use by governments should serve the public interest

The fundamental mission of governments, public institutions and public officials is to serve the public interest. The use of data by governments, public sector organisations and public officials should aim at contributing to public integrity[9] and delivering benefits for society. For this purpose, governments, public institutions and public officials should always prioritise the public interest over ill-intentioned or narrow private interests, and take into account the legitimate interests of stakeholders such as individuals, communities, and the private sector to maximise the benefits of data access, sharing and use for society as a whole.

Data use by governments should deliver public good

Observing cross-cutting values such as democracy, legitimacy, fairness, inclusion, transparency and openness is the sine qua non of ethical data use. Data use by governments, including the decisions and actions that derive from it, should prevent, avoid, or at the very least limit intentional harm to individuals, collectives and society as a manifestation of the principle of non-maleficence. To achieve this, governments should ensure the ethical management of data, including that of individuals and communities, throughout the data value cycle, while strengthening democratic institutions and the rule of law (e.g. in terms of data protection and personal privacy). This would help increase governments' legitimacy in the processing and use of data and deliver human-centred policies and services. Data use by governments should not lead to or perpetuate discrimination. It should instead promote inclusion, respect diversity, and ensure that individuals (e.g. citizens, residents) and collectives are treated and benefit equally from the outcomes that a data-driven public sector[10] aims to deliver.

ADDITIONAL CONSIDERATIONS

In developing the Good Practice Principles, the Thematic Group on Data-driven Public Sector noted that the following issues could be explored further by the E-leaders:

- **The collective and community nature of data governance:** In some national instances and contexts, the ethical use of data might imply acknowledging the collective and community aspects of data governance. This can include recognising the right levels of sovereignty and ownership of groups with specific rights (e.g. indigenous communities[11], neighbourhoods, stakeholder groups). Both individuals (including citizens and residents) and communities should be granted agency and self-determination over their data equally, including the definition of ad hoc data governance arrangements and the provision of the tools needed for this purpose. In this context, governments' stewardship of collectively-owned data should be duly acknowledged.
- **The environmental implications of data infrastructure:** Governments should take action to address the potential environmental impact of digital and data infrastructure. This includes, for instance, reducing their carbon footprint (e.g. avoiding the proliferation of unnecessary, redundant or overlapping data infrastructure such as data centres) and investing in clean and renewable energy infrastructure[12].
- **Abuse in the use of data during electoral campaigns:** The ethical use of data by politicians running for office reinforces trust in government and political parties, especially when these politicians are elected. The risks of using data together with tools for data analytics and behavioural insights during electoral campaigns call for further social and government oversight. This in order to help ensure that data and tools are not used to advance unethical goals such as misinformation or social manipulation, ensure the primacy of human autonomy and dignity, and sustain or increase trust in the election process and the legitimacy of the elected government.

The Good Practice Principles for Data Ethics in the Public Sector

MANAGE DATA WITH INTEGRITY

Data ethics is holistic. Public officials should always ensure trustworthy data management across the different stages of the data value cycle, which include, but are not limited to, data generation, collection, selection, curation, storage, disposal, access, sharing, and use. This is to maintain and strengthen public trust. For this purpose, public officials should:

- Not abuse their position, including the data at their disposal. Public trust in how the government manages data makes public officials data stewards by default. Public officials should act accordingly regardless of their position and role.
- Not access, share and use data for personal profit or for any goals that do not serve the public interest or undermine human rights.
- Manage data in accordance with applicable hard and soft regulatory instruments, including legislation, guidelines, formal recommendations, codes of conduct, self-assessment tools, integrity codes, and national and international standards.

BE AWARE OF AND OBSERVE RELEVANT GOVERNMENT-WIDE ARRANGEMENTS FOR TRUSTWORTHY DATA ACCESS, SHARING AND USE

It is the responsibility of public officials to be aware of and build knowledge in the specific governance arrangements, mechanisms and tools framing data access, sharing and use, to ensure they are respected, applied and used. For this purpose, public officials should:

- Be aware of, and respect, the different levels of ex-ante and ex-post responsibility and accountability proper to their role and context (e.g. policy area of work).
- Be aware of, and observe, the possibilities and limitations of data use and re-use and their expected behaviours, as defined by applicable rules listed in instruments such as policy, legislation, regulation, codes of conduct, and integrity codes.
- Be aware of, and acknowledge, the potential sanctions they might be subject to as a result of intended or unintended data abuse and mismanagement, and the resulting harm inflicted on individuals, collectives and society.
- Be aware of and, when needed, reach out to the authorities responsible for setting data governance rules and providing advice on the ethical implications of data access, sharing and use (e.g. organisational ethics commissions, high-level ethical policy advisory bodies, data protection bodies).
- Provide, have access to and/or make use of training opportunities to build, increase and share knowledge on the available data governance arrangements, mechanisms and tools supporting the ethical use of data, and develop the skills and certifications needed for this purpose.

INCORPORATE DATA ETHICAL CONSIDERATIONS INTO GOVERNMENTAL, ORGANISATIONAL AND PUBLIC SECTOR DECISION-MAKING PROCESSES

Public officials should recognise and take the appropriate measures to mitigate ethical risk at different levels so that this type of non-binding guidance leads to real impact. For this purpose, and depending on their position and level of responsibility, public officials might consider the incorporation of data ethical considerations as part of, or as preconditions for, inter alia:

- The generation of public sector data and decisions on data collection. Acknowledge that human bias or incomplete data can have a negative impact on i) the data inputs that inform policies and service design and delivery, and ii) the data outputs policies and services produce. This can lead to unintended outcomes such as discriminatory decisions or a partial view of problems.
- The planning and funding of public sector digital and data projects, including from the early stages of project design through to approval. Reinforce data ethics in the range of policy levers available to governments (e.g. revision and approval of digital projects by digital government bodies, conditional funding to enforce compliance with digital and technology standards) as well as in the accountability mechanisms public officials are subject to.
- The procurement and commissioning of public projects, in particular those involving the processing of personal data[13], personal sensitive data[14], or community data. Integrate data ethical considerations into contractual agreements or the terms and conditions of partnerships with third parties or external actors (e.g. as data use or sharing agreements, ethical checks, data management rules),

which would be collecting or using data on behalf of or in connection with their work for government.
- The processing of personal, personal sensitive or community data by third parties in the context of public-private partnerships should be transparent. It should comply with and adhere to applicable policy and legislation and with those rules and practices on data management supporting the ethical use of data in the public sector.

MONITOR AND RETAIN CONTROL OVER DATA INPUTS, IN PARTICULAR THOSE USED TO INFORM THE DEVELOPMENT AND TRAINING OF AI SYSTEMS, AND ADOPT A RISK-BASED APPROACH TO THE AUTOMATION OF DECISIONS.

The use of data and of AI systems based on data brings with it ethical responsibilities. Public officials[15] (e.g. those acting as data end-users or AI actors[16] in the public sector in connection with their public function) should retain control over the data they access, share and use to train AI systems. Also, public officials should not outsource to machines decisions that require unique insight into the human condition and have an adverse impact on human rights, democracy or the rule of law. For this purpose, and depending on their position and level of responsibility, public officials should:

- Monitor and control the quality, suitability and impartiality of data inputs (including large-scale datasets) by defining and deploying data management rules and practices (e.g. data documentation and validation) and creating an evidence trail to enable assessments of the trustworthiness of data and examinations of its provenance.

- Where appropriate, oversee the decisions made using or supported by an AI system. The level of human oversight should be congruent with the level of risk of the AI system to an individual, group, or business.
- Make sure any decisions that require unique human insight into the specific individual, social and economic context of impacted individuals or groups do not rely solely on automated processes. This includes decisions that could have an impact on individual, community and/or societal well-being and the public good (e.g. decisions determining access to or eligibility for public services).
- Establish frameworks or criteria to decide on and guide AI risk assessments, including to assess the sources and quality of data inputs used to train AI models, and define a timely and formal process to allow relevant parties to challenge the use or output of an AI system.
- Where needed, look for internal and external expertise to better understand the outputs (e.g. predictions, recommendations or decisions) delivered by AI systems to inform their final determinations. Be transparent, open and clear about data inputs and machine and/or human processes (e.g. criteria) that led to these final determinations.

BE SPECIFIC ABOUT THE PURPOSE OF DATA USE, ESPECIALLY IN THE CASE OF PERSONAL DATA

In designing digital and data-driven projects and initiatives, public officials should consider whether data needs to be collected, accessed, shared or used in the first place[17]. For this purpose, and depending on their position and level of responsibility, public officials should

- Ensure that the project proposition clearly articulates the purpose and legitimate interest that justifies the reason why data collection, access, sharing or use is needed.
- Make sure data is fit for purpose, representative of corresponding phenomena, and democratic in terms of the legitimacy and impact of the proposed use. Working on data quality dimensions such as completeness, comprehensiveness, consistency and accuracy[18] can help ensure data integrity and maximize the value of data to the purpose or problem it aims to solve.
- Be user-driven[19] and place users' needs and their concerns at the core of project design, implementation and monitoring. If the user need or the problem to be solved (e.g. the design and delivery of a digital service and deci-

sions on the access to public services) requires collecting or processing personal data, personal sensitive data, or community data, relevant stakeholders[20] or their representatives should be informed and their approval secured.
- Discern the practical implications of primary versus secondary data use while complying with applicable data regimes. Re-using data in a way that differs from the purpose for which it was initially collected, including data use by other actors besides the original user, might require defining and agreeing upon specific conditions or additional data management rules to maintain trust.

DEFINE BOUNDARIES FOR DATA COLLECTION, ACCESS, SHARING AND USE

It should be ensured that data governance and decision-making processes promote a balanced approach to data collection and use by weighing relevant trade-offs and societal costs and benefits; and assessing constraints, risks and rules surrounding data sharing, collection and use (including accredited access). In adopting norms such as data minimisation and proportionality, and considering using data aggregation and encryption tools, and depending on their position and level of responsibility, public officials should:

- Acknowledge that the type and use-context of data, not the technology nor the tool used for its processing, determine the relevant principles, rules and norms bearing on its use. In the case of personal data the norm of minimal data collection should be upheld. This could contrast with technologies like AI, which require significantly larger volumes of data ('big data') to evolve and function. In this light, exploring alternatives that might help to limit the collection[21] and use of personal data, personal sensitive data, or community data at the outset can also help in managing projects and reducing the number of rules to be complied with.
- Develop standardised decision-making instruments such as self-assessment/self-reflection tools, risk-maps and check-ups to gauge and justify the purpose and value of data use, including in relation to applicable rules. If feasible, the execution of these assessments should be mandatory for those projects involving the processing of personal data, personal sensitive data, or community data.
- Define rules for data management (e.g. data disposition and data retention, automatic data deletion), in particular in the context of projects collecting personal data, personal sensitive data, or community data. Ensure that such rules are clearly communicated to public officials in a timely manner. These rules should also be communicated to stakeholders or their representatives where appropriate, and framed by robust and efficient mechanisms to support their enforcement.
- Establish an agile[22] culture so that public servants are empowered to foster public sector exploratory innovation in a safe and controlled environment, test possible approaches, and experiment safely with data. Public sector agility[23] enables public servants to react promptly, iterate on approaches to the use of data and decide to change course in advance when risks are identified, objectives are not met or results are not achieved[24].

BE CLEAR, INCLUSIVE AND OPEN

Government openness[25] and public communication[26] are key to inform and engage relevant stakeholders or their representatives in an inclusive process of social dialogue around the ethical use of data in the public sector. Governments should be open about how data is being used, for what purpose, and by whom[27]. In this light, clarity[28] plays a key role in making sure that the recipient (e.g. data subject) understands the message. Also, data literacy goes beyond skills for data use for it can help data subjects to better understand the ethical implications of data use, including theirs. For this purpose, and depending on their position and level of responsibility, public officials should:

- Ensure the availability of multi-faceted and diverse teams working on or collaborating around specific projects. Diversity in the workplace can help to mitigate biases by offering multiple perspectives on a policy issue and fostering inclusive and informed decisions in terms of the data informing or resulting from a project (e.g. selection of data sources, data availability issues, data access restrictions, data's reflection of reality).
- Publish data governance and management policies, practices, and procedures, especially around the use of personal data.
- Engage in social dialogue with relevant actors inside and outside the public sector. These include actors whose data is being used, or their representatives, and secon-

dary stakeholders who can be affected or harmed by data use. Multi-stakeholder and multi-faceted approaches can help in identifying risks, defining boundaries and channeling actions prior, during and after the deployment of projects, policies and decisions involving the access to, sharing and use of data.

- Communicate to relevant stakeholders, or their representatives, in a clear and understandable way about the role of data (e.g. expected benefits and trade-offs), and its primary purpose – including in the context of training algorithms. Intention and use beyond the original purpose and the impact of not consenting to data use should also be communicated (e.g. delays due to slower decision-making procedures to grant access to or deliver public services).
- Acknowledge the social context, including factors such as the presence of indigenous communities and native non-official languages to foster inclusion.
- Educate relevant stakeholders (e.g. data subjects and their representatives, and those from vulnerable, underrepresented, or marginalised groups in society) on data governance, including its meaning and implications for them. Confront scenarios in which only privileged and educated segments of the population have a voice and say in how their data is being used. This includes the capacity to contest certain uses of data.

PUBLISH OPEN DATA AND SOURCE CODE

Open data and open source code help reap socio-economic benefits and foster citizen engagement and innovation while securing the transparency, accountability, and public scrutiny of governments' decisions and policy outcomes. For this purpose, and depending on their position and level of responsibility, public officials should:

- Promote fair data ecosystems through open government data (OGD) policies[29]. By sharing public sector data as open data, public officials grant unrestricted access to valuable data sources and help to ensure that the benefits of data are equitably distributed in society, contribute to the public good, and create public value.
- Open disaggregated and granular data in accordance with applicable privacy, security, and ownership requirements. When shared under the right conditions (e.g. as anonymised data), open data can help identify social, economic, and other inequalities, address data gaps, and promote evidenced-informed decision-making.
- Connect open government data initiatives to broader data governance and management efforts in the public sector. This is to ensure that open data aligns to efforts aiming at the mitigation of biases affecting the generation or collection of data by public sector organisations (e.g. open data is not fully representative of a phenomenon).
- Make source code openly available for public scrutiny and audit, in particular when personal data, personal sensitive data, or community data is processed as part of digital government projects (e.g. contact tracing apps).

BROADEN INDIVIDUALS' AND COLLECTIVES' CONTROL OVER THEIR DATA

Upon being informed about how and with whom personal and collectively owned data is shared, individuals (including citizens and residents) and communities should be given decision-making power to exercise autonomy, control, and agency over their data, and to freely give or withdraw consent to its use. For this purpose, and depending on their position and level of responsibility, public officials should:

- Offer data subjects or their representatives the possibility and tools to opt-in to and opt-out of specific data uses (e.g. design of digital services, AI training), and delete personal data records if they so choose (e.g. as features of government-operated apps).
- Draw upon and put in place the right transparency and digital tools (such as digital identity and citizens' folders) to this purpose. More advanced data access and sharing arrangements (e.g. data trusts, data fiduciaries) can help in this respect depending on the complexity of the data and the purpose of its use.
- Design and deploy the needed tools (or build upon existing mechanisms such as Freedom of Information requests) to enable individuals to request information from public sector organisations on their data holdings, namely on the data they hold about the requester (data subject).
- Avoid creating multiple copies of the same data, in particular of personal data and personal sensitive data, and use shared data infrastructures for data storage and management. Better control over sensitive data assets helps implement digital government principles such as once only and ensure that, when needed, data subjects receive a timely and accurate view of the data public organisations hold about them.

BE ACCOUNTABLE AND PROACTIVE IN MANAGING RISKS

Governments are increasingly equipped to anticipate and proactively address public concerns regarding the collection, access, sharing and use of personal, personal sensitive data or community data. Accountable data use involves ensuring that i) data users comply with all applicable policy, legislative, and regulatory requirements by design; ii) clear and common data management rules are in place to support the fair and trustworthy access, sharing and use of data; iii) data governance structures are available to provide advice, intervene or correct actions; and iv) relevant bodies such as parliaments and judicial bodies can intervene when needed. For this purpose, and depending on their position and level of responsibility, public officials should:

- Build-in procedures to systematically address potential deviations. This is especially important when data is falsified or misrepresented in order to satisfy partial interests (e.g. fake data).
- For those projects already in place, encourage data users to self-evaluate the data they are using or reusing, and the ways that data has been generated or collected (especially in the case of data obtained from external or third parties). Retroactive assessments can help in aligning existing projects with new rules, checks and controls, which were not in place prior.
- Promote peer-to-peer assessments among public officials when developing and implementing data-driven projects. This can facilitate the exchange of views and expertise and the identification of otherwise ignored implications and risks related to the handling of data.
- Perform regular and random data interventions (data audits). These oversight actions should not only assess data quality, including compliance with standards, but also evaluate whether data is fit for purpose and ensure that its use is proportionate and legitimate. Interventions should also examine whether data processing complies with best practices and established rules, and monitor intended versus actual outcomes.
- Enable actors such as internal and external auditors to perform data auditing tasks. Standardised evaluation instruments (e.g. self-assessment/self-reflection tools, risk-maps and check-ups) and well-documented data strategies (e.g. data handling practices, training records, or data residency rules for sensitive data) create audit trails for these actors to perform their role effectively.
- Develop or use available internal and external communication channels for public officials and data subjects to submit inquiries or remarks, raise identified or imminent risks, contest decisions, or report data collection or processing errors, data misuse, and unintended outcomes. Be responsive to such inputs and ensure that any concerns are addressed effectively.
- Create safe havens for reporting data misuse, unintended negative outcomes and early warnings, and protect whistle-blowers reporting wrongdoing.
- Be transparent and accountable when events such as data misuse and data leaks occur. This can help maintain the community's trust in governments as data stewards and in the capacity of public officials to act accordingly. Projects should be stopped if necessary to prevent further damage and harm.

Annex A. Methodology for development

The Good Practice Principles for Data Ethics in the Public Sector were drawn together by the Thematic Group on Data-driven Public Sector under the aegis of the OECD Working Party of Senior Digital Government Officials (E-leaders).

The OECD Working Party of Senior Digital Government Officials (E-Leaders), a subsidiary body of the Public Governance Committee, together with Thematic Groups working under its aegis, delve into topical issues of relevance to digital government, including data governance in the public sector. This involves supporting projects and pilots that are anticipated or already underway across OECD Member and certain non-Member countries in order to identify key policy actions for governments in priority areas identified by Thematic Group delegates. Much of the work of the Thematic Groups has sought to respond to demands for specific guidance to support the implementation of the OECD Recommendation on Digital Government Strategies in priority areas.

The Thematic Group on Data-driven Public Sector emerged in 2018 from the former Thematic Group on Personal Data Ownership and Transparency. The activities of and rationale behind the Thematic Group on Data-driven Public Sector draw upon the decisions taken by E-Leaders following the 2015 meeting of the Working Party held in Tokyo, Japan. After the meeting, government technology leaders from OECD Member and non-Member countries agreed upon the need to foster a data-driven public sector, including through big data and open data. They also discussed the urgency of maintaining and building public trust through increased transparency, while ensuring the security and privacy of personal data (OECD, 2015)[30].

Between 2018 and 2020, the activities of the Thematic Group have cycled through various stages due to changes in the Group's focus of the work in 2018 and 2019 that were agreed upon by participating countries. The Netherlands has led the activities of the Thematic Group since 2018[31]. The main goals of the Thematic Group on Data-driven Public Sector in 2018 focused on:

- sharing experiences of how governments are using data-driven methods and techniques to improve their organisations, policy processes and decisions, service delivery, and public sectors in general.
- sharing experiences of how governments are implementing these data-driven methods and techniques.
- sharing and developing guidelines on the use of data (e.g. how to apply the principle of transparency to the ways in which data is used in the public sector).

The results of the Thematic Group's activities in 2018 were presented and discussed during the 2018 Meeting of the E-Leaders held in Seoul, South Korea on 30-31 October, 2018. The session sought to shed light on how governments could leverage data as a strategic asset to boost digital maturity in the public sector and, as a result, develop the conditions and capabilities needed for sustainable and inclusive policies and services. These capabilities could, for example, support the ethical development and use of AI for public value co-creation.

The discussions during the 2018 Meeting of the E-Leaders led to a refocusing of the activities of the Thematic Group for Q4 of 2018 and the first half of 2019[32], thus aiming at developing guidelines, frameworks, and tools that could help frame countries' decisions and actions regarding the implementation of data initiatives with an ethical approach. As a result, participating countries agreed on working towards the development of a common set of guidelines and/or principles for data ethics within the public sector.

A first version of Data Ethics Guidelines was presented during the 5th Meeting of the OECD Expert Group on Open Government Data (6-7 June, 2019), and during the 2019 Meeting of the E-Leaders held in Brussels, Belgium on 19-20 September, 2019.

In 2020/21, the activities of the Thematic Group on Data-driven Public Sector aimed at refining and finalising the draft Good Practice Principles, including:

- A series of co-ordination calls with delegates participating in the Thematic Group;
- The collection of comments and insights from different OECD units working on open government, public sector innovation, evidence-based policy making, public sector integrity, procurement and infrastructure, and regulation;
- A targeted consultation to collect views from international partners;

- A consultation round with delegates from the OECD Expert Group on Open Government Data; and
- The collection of comments from other relevant Directorates within the OECD.

The development of the draft Good Practice Principles for Data Ethics in the Public Sector presented in this document benefited from a gap analysis performed by the OECD Secretariat. The gap analysis explored the values and issues raised in existing ethical principles addressing the use of data. This exercise allowed for identifying and highlighting aspects that should be taken into consideration to promote greater impact and value of the proposed Good Practice Principles for Data Ethics in the Public Sector for OECD governments, policy and decision makers, politicians, project managers and data practitioners in the public sector.

NOTES

[1] For a definition of data ethics see for instance the UK Data Ethics framework at https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-glossary-and-methodology which provides a definition based on those from Luciano Floridi, Mariarosaria Taddeo (2016) 'What is data ethics?'; Pernille Tranberg, Gry Hasselbalch, Birgitte Kofod Olsen & Catrine Søndergaard Byrne (2018) 'Data Ethics. Principles and Guidelines for Companies, Authorities & Organisations. See also the definition provided by the Open Data Institute at https://theodi.org/article/data-ethics-canvas/#1562602644259-1d65b099-ea7b , and the work of the Alan Turing Institute on data ethics at https://www.turing.ac.uk/research/data-ethics for more information on the topic.

[2] For the purpose of these Good Practice Principles a data regime is understood as the global, regional, national and/or local norms and regulations on data. A data regime can refer to written legislation and regulations addressing issues such as data access and sharing within the public sector, across sectors (e.g. G2B), data protection, or open data. Therefore, data ethics aim at agreeing upon a common set of shared values-based rules to guide customary data practices while complementing formal written norms and regulations.

[3] For more information see: OECD (2019), State of the art in the use of emerging technologies in the public sector , OECD Working Papers on Public Governance No. 31, September 2019. Available at: https://www.oecd.org/gov/digital-government/working-paper-the-use-of-emerging-technologies-in-the-public-sector.htm

[4] For more information see: http://www.oecd.org/gov/the-path-to-becoming-a-data-driven-public-sector-059814a7-en.htm

[5] The OECD Recommendation of the Council on Artificial Intelligence defines AI systems as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy". For more information see: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[6] The OECD Privacy Guidelines are an integral part of the Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data [OECD/LEGAL/0188]. See also http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

[7] For more information see: OECD (2019), Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies, OECD Publishing, Paris, https://doi.org/10.1787/276aaca8-en.

[8] For more information see: OECD (2019), The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/059814a7-en.

[9] The OECD Recommendation on Public Integrity 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". For more information see: https://www.oecd.org/gov/ethics/OECD-Recommendation-Public-Integrity.pdf

[10] For more information see: OECD (2019), The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/059814a7-en.

[11] See for instance the case of Canada at https://www.afn.ca/uploads/files/nihbforum/info_and_privacy_doc-ocap.pdf & international efforts such as the CARE Principles for Indigenous Data Governance at https://www.gida-global.org/care

[12] For more information, see for instance: Lucivero, F. Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives. Sci Eng Ethics 26, 1009–1030 (2020). https://doi.org/10.1007/s11948-019-00171-7 & IEA (2019), Data centres and energy – from global headlines to local headaches?, IEA, Paris https://www.iea.org/commentaries/data-centres-and-energy-from-global-headlines-to-local-headaches

[13] The OECD Privacy Guidelines define personal data as: "any information relating to an identified or identifiable individual (data subject)". For more information see: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

[14] For more details see: https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en

[15] The OECD defines public officials as "people who hold a legislative, administrative or judicial office (either appointed or elected); any person exercising a public function, including for a public agency or a public enterprises (e.g. a state owned enterprise); any official or agent of a public international organisation". For more information see: https://stats.oecd.org/glossary/detail.asp?ID=7252#:~:text=These%20include%20people%20who%20hold,of%20a%20public%20international%20organisation.

[16] The OECD Recommendation of the Council on Artificial Intelligence defines AI actors as "those who play an active role in the AI system lifecycle, including organisations and individuals that deploy or operate AI". For more information see: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[17] The OECD Privacy Guidelines introduced the Purpose Specification Principle, which states that "the purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose". For more information see: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

[18] For more information, see From data quality to data qualities. Chapter 3: Leveraging accessibility through high-quality open data. Page 89-90, in OECD (2018), Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264305847-en.

[19] For more information on the application of user-driven approaches in the context of digital government see: OECD (2020), "The OECD Digital Government Policy Framework: Six dimensions of a Digital Government", OECD Public Governance Policy Papers, No. 02, OECD Publishing, Paris, https://doi.org/10.1787/f64fed2a-en.

[20] The OECD Recommendation of the Council on Open Government defines stakeholders as "any interested and/or affected party, including: individuals, regardless of their age, gender, sexual orientation, religious and political affiliations; and institutions and organisations, whether governmental or non-governmental, from civil society, academia, the media or the private sector". For more information see: https://www.oecd.org/gov/Recommendation-Open-Government-Approved-Council-141217.pdf

[21] The OECD Privacy Guidelines introduced the Collection Limitation Principle, which states that "there should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject". For more information see: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

[22] For more information on agility in the public sector see for instance https://www.gov.uk/service-manual/agile-delivery/agile-government-services-introduction

[23] See for instance the OECD Policy Framework on Sound Public Governance and the role of public sector innovation and agility in this context at: https://www.oecd.org/governance/pcsd/OECD_Policy_Framework_on_Sound_Public_Governance_Highlights%20Brochure_EN.pdf

[24] For more information on exploration, iteration and testing in the context of public sector innovation see the OECD Declaration on Public Sector Innovation at https://oecd-opsi.org/wp-content/uploads/2018/11/OECD-Declaration-on-Public-Sector-Innovation-English.pdf

[25] The OECD Recommendation of the Council on Open Government defines Open Government as "a culture of governance that promotes the principles of transparency, integrity, accountability and stakeholder participation in support of democracy and inclusive growth". For more information see: https://www.oecd.org/gov/Recommendation-Open-Government-Approved-Council-141217.pdf

[26] Public communication "is understood as any communication activity or initiative led by public institutions for the public good. It is different from political communication, which is linked to the political debate, elections, or individual political figures and parties. Public communication activities can include the provision of information, as well as consultation and dialogue with stakeholders". For more information on the OECD work on public communication see: http://www.oecd.org/coronavirus/policy-responses/transparency-communication-and-trust-bef7ad6e/#back-endnotea0z3 & https://www.oecd.org/gov/open-government/oecd-international-report-on-public-communication.htm

[27] The OECD Privacy Guidelines introduced the Openness Principle, which states that "there should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller." For more information see: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

[28] According the European General Data Protection Regulation, "the principle of transparency requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand and that clear and plain language be used". For more information see: https://gdpr.eu/tag/gdpr/

[29] For more information on the OECD work on open government data see: https://www.oecd.org/digital/digital-government/open-government-data.htm

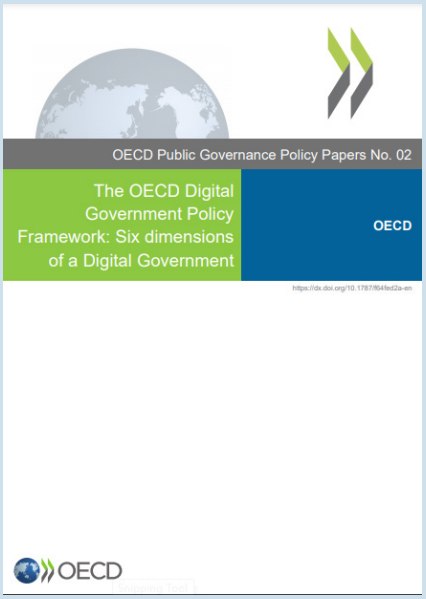
[30] OECD E-Leaders Meeting 2015: Communiqué. http://www.oecd.org/governance/eleaders/oecd-e-leaders-meeting-2015-communique.htm

[31] In 2018, the Thematic Group benefited from the participation of delegates from Argentina, Colombia, Chile, Egypt, Estonia, Latvia, Slovenia, the United Kingdom, and Uruguay.

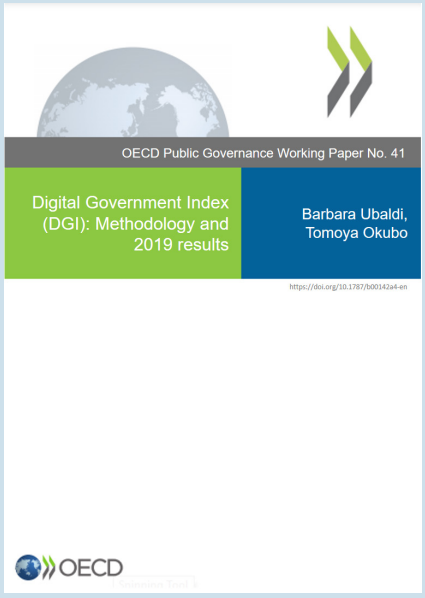
[32] In 2019, members of the Thematic Group included Australia, Argentina, Belgium, Brazil, Canada, Chile, Colombia, Denmark, Egypt, Estonia, Finland, Greece, Israel, Italy, Ireland, Latvia, Lithuania, Luxembourg, Poland, Portugal, Panama, Slovenia, South Korea, Spain, Sweden, the United Kingdom, the Netherlands, and Uruguay.

OECD Digital Government Index: 2019 results
https://oe.cd/dgi2019

2014 Recommendation of the Council on Digital Government
https://oe.cd/digitalgovrecommendation

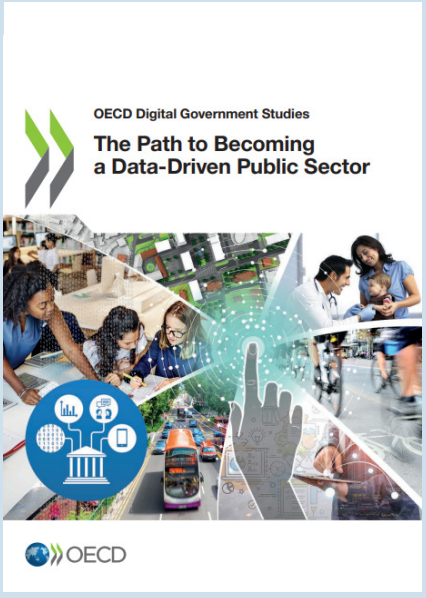


OECD Digital Government Policy Framework
https://oe.cd/il/diggovframework



OECD Digital Government Index: Methodology and 2019 results
https://oe.cd/il/dgimethodology

OECD Open, Useful and Re-usable (OUR) Data Index: 2019 results
https://oe.cd/open-data-2019



2019 The Path to Becoming a Data-Driven Public Sector report
https://oe.cd/il/ddps


digital
Making the transformation work
for growth and well-being


OECD