Title: A framework for regulating AI in the public sector

Adam Panagiotopoulos and Rowena Rodrigues (corresponding author), Trilateral Research Ltd.
Email: rowena.rodrigues@trilateralresearch.com

Abstract: Recent technological and regulatory developments in the public sector across the globe signal a radical transition; the traditionally technologically conservative paper-based public sector is transitioning to a digitalised, tech-savvy, advanced and inter-connected system of public services, assimilating market features. Public services and infrastructure are being digitalised; artificial intelligence (AI) and other innovative technologies are being deployed by governmental and local services. There is an increasing trend for governments to set the AI-driven digital transformation of the public sector as a key priority in their regulatory and political agendas. This chapter presents the implications of the digitalised public sector, considering both the positive and negative repercussions of AI in the sector. It also analyses the trends and challenges of regulating AI in the public sector, taking into account shared attributes in regional and international regulatory frameworks and strategies. Based on these findings, we propose a framework for regulating AI in the public sector bringing together what have been identified as critical elements to ensure that the deployment and use of AI in this sector do not prevent States from performing their obligations and duties towards society and individuals. The framework covers three core aspects: (i) Legal (ii) Ethical and (iii) Societal. This framework will help advance the governance of the use of AI and provide insights into the critical elements that need to be present and further developed. This framework aims at laying down the conditions for the full potential of AI deployment and use in the public sector in a manner that meets societal expectations, is in line with fundamental values, rights and freedoms and is underpinned with ethical principles.

Keywords: artificial intelligence, AI regulation, public sector, governance, law, assessment

I.   Introduction

The concept of the public sector is very broad and depends on the legal, historical, and constitutional traditions and practices of States. In general, it can include central government, state government, local government, social security funds, public corporations and non-market non-profit institutions controlled and financed by government units.[1] Overall, the functions of the government cover the areas of general public services, defence, public order, economic affairs, environmental protection, housing and community amenities, health, recreation, culture and religion, education and social protection.[2] The structural and functional reorganisation of the public sector and its services through digital drivers has been a key item on the governmental agenda for over the past three decades since the early computing developments in the 1950s.[3] Since the 1960s, governments have initiated the

---

[1] OECD, Glossary Of Statistical Terms - Government Sector Definition, https://stats.oecd.org/glossary/detail.asp?ID=1139.

[2] Eurostat, Glossary: Classification of the functions of government (COFOG), https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Classification_of_the_functions_of_government_(COFOG).

[3] van Ooijen, C., B. Ubaldi and B. Welby (2019), "A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance", OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris, https://doi.org/10.1787/09ab162c-en. p 9.

digital transformation of the public sector (e.g., census studies, health records). This transformation spans the design and delivery of public services, including enforcement and control, and builds on gradual transformation waves from the analogue government towards the digital government.[4]

In particular, European public administrations are currently progressing through the wave of e-government, where information and communication technologies enable "borderless, personalised, user-friendly, end-to-end digital public services to all citizens and businesses in the EU".[5] At the European level, overall, there is a widespread and intense digitalisation of public services throughout the 27 EU Member States with establishing and promoting e-government practices and ICT-enabled innovation. In general, both internal administrative systems and public-facing interfaces and interactions are being digitalised, whereas the European public sector is setting up innovation and artificial intelligence (AI) hubs and centres.[6] The European eGovernment Action Plan 2016-2020 laid down seven guiding principles for the digitalisation of European public administrations, including the premise of Digital by Default.[7] As part of this ongoing digitalisation, emerging technologies also support digital sovereignty and governance in the internal and external relationships of States. AI is leading the digital transformation of the public sector and constitutes a core driver in the national strategies of digital transformation.[8]

These recent technological and regulatory developments in the public sector across the world signal a radical transition; the traditionally technologically conservative paper-based public sector is transitioning to a digitalised, tech-savvy, advanced and inter-connected system of public services, assimilating market features.[9] Governments are modernising their systems focusing on the efficiency and effectiveness of the delivery of public services. Services and infrastructure are being digitalised, whereas AI and other innovative technologies are being deployed by governmental and local services. This is also mirrored at the citizen level where individuals are being both actively and indirectly datafied as a condition for effective access to public services. The deployment and application of AI solutions are no longer the privileges of the Big Tech and start-ups, but States are also taking the lead. In fact, there is an increasing trend for governments to set the AI-driven digital transformation of the

---

[4] See van Ooijen, C., B. Ubaldi and B. Welby (2019), "A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance", OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris, p. 9, https://doi.org/10.1787/09ab162c-en, and A comparative overview of public administration characteristics and performance in EU28 2017 p. 46 -48, https://op.europa.eu/en/publication-detail/-/publication/3e89d981-48fc-11e8-be1d-01aa75ed71a1/language-en.

[5] Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions, EU eGovernment Action Plan 2016-2020 Accelerating the digital transformation of government, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016DC0179.

[6] See indicatively A multi-dimensional framework to evaluate the innovation potential of digital public services - A step towards building an Innovative Public Services Observatory in the EU, 2020 p. 19-21.

[7] Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions, EU eGovernment Action Plan 2016-2020 Accelerating the digital transformation of government, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016DC0179.

[8] Misuraca, G., Barcevičius, E. and Codagnone, C., Exploring Digital Government Transformation in the EU - Understanding public sector innovation in a data-driven society, 2020, https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/exploring-digital-government-transformation-eu-understanding-public-sector-innovation-data

[9] Painter, Martin et al (2010), Tradition and Public Administration, Palgrave Macmillan.

public sector as a key priority in their regulatory and political agendas. Yet, whereas a large amount of research has been conducted into the technical, ethical and legal requirements and implications of AI-driven systems, the question of the regulation of these systems in the public sector has not been given adequate attention. Research demonstrates that the public sector often implements and invests in AI technologies under opaque legal conditions and inadequate organisational, policy and technical frameworks.[10]

This chapter discusses the benefits, challenges and pitfalls of the AI-driven digitalisation of the public sector and current regulatory approaches to AI in the public sector. The structure of this chapter is as follows: first, we look at the implications of the digitalised public sector, considering both positive and negative repercussions of the AI applications in the public sector. We then examine the trends and challenges of regulating AI in the public sector, taking into account shared attributes in regional and international regulatory frameworks and strategies. Based on these findings, we propose a framework for regulating AI in the public sector bringing together what we have identified as critical elements. This is significant because, as noted, "a strong and coherent regulatory framework for AI" is still very much a work in progress.[11] This framework draws on assessment questions for regulating AI in the public sector, bringing together insights and findings on the critical elements that underpin the deployment and use of AI in the public sector. It covers three core aspects: (i) Legal (ii) Ethical and (iii) Societal aspects and will help advance the governance of the use of AI in the public sector and provide insights into the critical elements that need to be present and further developed. This framework aims at laying down the conditions for the full potential of AI deployment and use in the public sector in a manner that meets societal expectations, is in line with fundamental values, rights and freedoms and is underpinned with ethical principles. This chapter was developed in 2020 based on research conducted in 2019 and prior to the publication (21 April 2021) of the Commission Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain legislative acts.[12] While we have referred to this Proposal it has not been part of the analysis, as the latter significantly preceded its release.

II.   Implications of the digitalised public sector

Despite the increasing AI omnipresence, the AI paradox lies in the difficulty to understand and delineate AI contrary to the easiness and eagerness of incrementally implementing AI across the public and private realm. A commonly accepted definition of AI systems refers to "software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital

---

[10] See indicatively Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute, p. 23-24, https://www.turing.ac.uk/research/publications/understanding-artificial-intelligence-ethics-and-safety and Big Brother Watch, Submission to the Centre for Data Ethics and Innovation: Bias in Algorithmic Decision-Making (Crime and Justice), June 2019, https://bigbrotherwatch.org.uk/wp-content/uploads/2019/06/Big-Brother-Watch-submission-to-the-Centre-for-Data-Ethics-and-Innovation-Bias-in-Algorithmic-Decision-Making-Crime-and-Justice-June-2019.pdf.

[11] See in the UK context, Committee on Standards in Public Life, Artificial Intelligence and Public Standards, 2020, p. 39, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF.

[12] COM/2021/206 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.

dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions."[13]

Individuals are already witnessing the digital transformation of States in their everyday interactions with the public system, whereas innovative technologies are being trialled and implemented across federal, central, and local administrations. States are rolling out AI-driven products and services in an effort to automate and digitalise their functions, which leads to structural and functional transformations of the core and peripheral public sector. Research into the uses of AI by States showcase that advanced States are one step ahead of the trial phase and are actively deploying AI systems.[14] Governments either proactively and willingly, or upon investigations, media exposure and Freedom of Information requests[15] share information about AI systems integrated into the public sector. Looking into the AI uses in the international public arena, it is clear that the level of AI engagement, dependency and maturity alongside the AI-covered cases, areas and public functions vary among the States. According to an analysis of the intelligence readiness of 194 governments to use AI in the delivery of public services, the Global North is best placed to implement AI systems.[16] Indeed, the Nordic Council of Ministers issued a declaration in 2017 to create a common area for cross-border digital services in the public sector,[17] followed by a declaration for enhanced AI in 2018.[18] It was also found that Singapore's public sector is most AI-savvy, followed by Western European governments as well as Canada, Australia, New Zealand, and four Asian economies in the top 20.[19] Despite this AI heterogeneity due to financial and technological resources and policy priorities, there is evidently an increase in the deployment of AI at the international, regional and national levels.

Transformative AI has contributed to the re-imagination of the traditional public structure as a digital state. The main AI solutions deployed in the public sector include virtual assistants, automatic handwriting recognition, speech recognition, natural language generation, natural language processing, computer vision, machine learning, deep learning, swarm intelligence and affective computation.[20] The AI solutions spread across various critical functions of States and tend to further

---

[13] High-Level Expert Group on Artificial Intelligence (AI HLEG), A definition of Artificial Intelligence: main capabilities and scientific disciplines, 2019, p.1, https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines.

[14] P. Misuraca, G., and van Noordt, C., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, https://joinup.ec.europa.eu/sites/default/files/document/2020-07/jrc120399_Misuraca-AI-Watch_Public-Services_30062020_DEF_0.pdf.

[15] See for example Bureau of Investigative Journalism (2019), Government Data Systems: The Bureau Investigates. Available at www.thebureauinvestigates.com/stories/2019-05-08/algorithms-government-it-systems.

[16] Oxford Insights, Government Artificial Intelligence Readiness Index, 2019, p. 5, https://ai4d.ai/wp-content/uploads/2019/05/ai-gov-readiness-report_v08.pdf.

[17] The Nordic-Baltic region: a digital frontrunner, 2017, https://www.norden.org/en/declaration/nordic-baltic-region-digital-frontrunner.

[18] AI in the Nordic-Baltic region, 2018, https://www.norden.org/en/declaration/ai-nordic-baltic-region

[19] Oxford Insights, Government Artificial Intelligence Readiness Index, 2019, https://ai4d.ai/wp-content/uploads/2019/05/ai-gov-readiness-report_v08.pdf.

[20] For a complete overview of the AI solutions deployed in the public sector and their definitions see OECD, Working Papers on Public Governance No. 31, State of the art in the use of emerging technologies in the public

expand to new uses and other public areas. Regarding the critical public sector areas, AI applications target human-machine interactions, mainly in the domain of health, transportation and security.[21] In particular, AI solutions are applied to the entire policy-making and service design process and AI applications are adopted across the spectrum of public sector services,[22] including education, energy[23], social investment, healthcare (such as patient treatments and optimisation of the everyday operation of hospitals)[24], housing and community relations, information society and digitalisation, justice[25], public order, law enforcement (such as predictive policing and crime hotspot analytics[26]), public administration modernisation and reform (such as classification and natural language processing to process unstructured data in online surveys[27] and virtual assistants to reply to individuals' questions), immigration and border control,[28] recreation, culture and religion, science,

---

sector, p. 73-74, https://www.oecd-ilibrary.org/docserver/932780bc-en.pdf?expires=1591370140&id=id&accname=guest&checksum=13388E68BC9300E63E066A21C3A31B07 . See also Deloitte Report, AI-augmented government -Using cognitive technologies to redesign public sector work AI-augmented government, 2017. https://www2.deloitte.com/content/dam/insights/us/articles/3832_AI-augmented-government/DUP_AI-augmented-government.pdf.

[21] OECD, Working Papers on Public Governance No. 31, State of the art in the use of emerging technologies in the public sector, p.4, https://www.oecd-ilibrary.org/docserver/932780bc-en.pdf?expires=1591370140&id=id&accname=guest&checksum=13388E68BC9300E63E066A21C3A31B07.

[22] See indicatively, Jamie Berryhill & Kévin Kok Heang & Rob Clogher & Keegan McBride, 2019, "Hello, World: Artificial intelligence and its use in the public sector," OECD Working Papers on Public Governance 36, OECD Publishing, p.72-88, https://www.oecd-ilibrary.org/governance/hello-world_726fd39d-en and S. Weslei et al, How and where is artificial intelligence in the public sector going? A literature review and research agenda Government Information Quarterly Volume 36, Issue 4, October 2019, p. 1-14.

[23] Government Digital Service and Office for Artificial Intelligence, Using data from electricity meters to predict energy consumption, 2019, https://www.gov.uk/government/case-studies/using-data-from-electricity-meters-to-predict-energy-consumption.

[24] See indicatively the Danish Government, National Strategy for Artificial Intelligence, 2019 p. 55, https://en.digst.dk/media/19337/305755_gb_version_final-a.pdf.

[25] Government Digital Service, Office for Artificial Intelligence, and Ministry of Justice, How the Ministry of Justice used AI to compare prison reports, 2019, https://www.gov.uk/government/case-studies/how-the-ministry-of-justice-used-ai-to-compare-prison-reports--2.

[26] UNICRI, Artificial Intelligence and Robotics for Law Enforcement, 2019, p. 5 https://www.europarl.europa.eu/cmsdata/196207/UNICRI%20-%20Artificial%20intelligence%20and%20robotics%20for%20law%20enforcement.pdf

[27] Government Digital Service and Office for Artificial Intelligence, How GDS used machine learning to make GOV.UK more accessible, https://www.gov.uk/government/case-studies/how-gds-used-machine-learning-to-make-govuk-more-accessible and Deloitte Report, Using AI to unleash the power of unstructured government data, 2018, https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/public-sector/lu-ai-unstructured-government-data.pdf.

[28] E.g., Big Data predictions about population movements, automated decision-making in immigration applications, AI detectors at borders. https://edri.org/the-human-rights-impacts-of-migration-control-technologies/.

research and innovation, taxation, duties and welfare payments, transport and infrastructure,[29] such as prediction of travel behaviour.[30]

In the European public sector, a review of AI adoption in public services in all 27 EU Member States, Norway, Switzerland and the UK, revealed that most AI solutions in the public sector support the provision of general public services or are deployed in communication and engagement activities.[31] Most AI solutions benefit central government functions, including general public services, economic affairs and health.[32] This study also found out that the level of AI adoption and use across European countries is highly diverse and heterogeneous "both in terms of the number of use cases identified, and of the different technologies regarded as AI and their purposes and functionalities."[33]

AI in the public sector revolutionises the design and delivery of public services.[34] In their capacity as AI developer, procurer, and applier, governments rely on AI tools to enable the public sector to process larger amounts of information and conduct forecasts and predictions for evidence-based decisions and policy-making.[35] In addition, the use of AI supports operational efficiency and relieves the government from time-consuming and burdensome tasks by automating physical and digital tasks. AI-driven systems support the rationalisation and monitoring of processes and improve the quality of public services.[36] This enables more cost-efficient and proportionate investments and spending of resources on high-level tasks. AI also personalises public services to the individual circumstances and conditions, which enhances the citizenship experience and engagement. AI in the European public sector is also relied upon "to predict needs, habits and behaviours of citizens, and consequently use

---

[29] Government Digital Service, Office for Artificial Intelligence, and Driver and Vehicle Standards Agency, How the Department for Transport used AI to improve MOT testing, 2019, https://www.gov.uk/government/case-studies/how-the-department-for-transport-used-ai-to-improve-mot-testing and Government Digital Service and Office for Artificial Intelligence, How a signalling company used AI to help trains run on time, 2019, https://www.gov.uk/government/case-studies/how-a-signalling-company-used-ai-to-help-trains-run-on-time.

[30] Norwegian Ministry of Local Government and Modernisation, National Strategy for Artificial Intelligence, p. 29, 2020, https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf.

[31] P. Misuraca, G., and van Noordt, C.., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, p. 81-82, https://joinup.ec.europa.eu/sites/default/files/document/2020-07/jrc120399_Misuraca-AI-Watch_Public-Services_30062020_DEF_0.pdf.

[32] Misuraca, G., and van Noordt, C., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, p. 16-21, https://joinup.ec.europa.eu/sites/default/files/document/2020-07/jrc120399_Misuraca-AI-Watch_Public-Services_30062020_DEF_0.pdf.

[33] Misuraca, G., and van Noordt, C.., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, p. 84, https://joinup.ec.europa.eu/sites/default/files/document/2020-07/jrc120399_Misuraca-AI-Watch_Public-Services_30062020_DEF_0.pdf.

[34] OECD, Working Papers on Public Governance No. 31 and A guide to using artificial intelligence in the public sector.

[35] Centre for Public Impact, Destination unknown: Exploring the impact of Artificial Intelligence on Government, 2017, p.27-32, https://resources.centreforpublicimpact.org/production/2017/09/Destination-Unknown-AI-and-government.pdf.

[36] Norwegian Ministry of Local Government and Modernisation, National Strategy for Artificial Intelligence, 2020, p. 53, https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf.

these predictions either for creating a more secure society or to deliver tailored services to more granular needs."[37]

The AI-driven digitalisation of the public sector comes with great benefits, as explained above, but as an innovative technology, poorly regulated in its infancy, it could challenge the way the States understand their role, functions and commitments to the public. Fundamental concepts that are commonly accepted among European public systems and support public accountability and democratic governance are being challenged and re-configured in the public discourse. Representative examples include the tenets of good administration, non-discrimination, due process, confidence in the public sector, the pursuit of the public good, transparency and public scrutiny, administrative procedural principles, proportionality, right to be heard, right to a judicial remedy, legality and legal certainty.[38] For example, the United Nations Special Rapporteur on extreme poverty and human rights commenting on the digital transformation of the UK state, as heralded in 2017, observed that "we are witnessing the gradual disappearance of the post-war British welfare state behind a webpage and an algorithm" and "in its place a digital welfare state is emerging".[39] This statement accentuates the challenges that digitalisation and AI-based innovation could inflict on the public sector in the absence of an appropriate governance framework.

In the digital administrative net, the users of algorithmic-driven public services are at risk of being subject to unfair and discriminatory treatment, where the AI tools replicate or reinforce the bias of the training data or its designers. Under a maximised approach for more data collection and matching across the public services to support Big Data in the public sector, "new state-citizen power dynamics are created as citizens/individuals become infinitely 'knowable'" and "emphasis shifts from causation to correlation and from prevention to pre-emption, prediction and probability".[40] In addition, interactive citizen-state interfaces, automated decision-making tools and systems could obscure who is accountable for the final output. Private actors are involved in the design and implementation of the digitalisation of the public sector while funded by the government, whereas private entities can also provide AI-driven services on behalf of the government impacting transparency and accountability.[41] Individuals are then confronted with a digital labyrinth, where they fail to hold the relevant parties accountable in the public or private sector.[42] This state of lack of access to redress

---

[37] Van Noordt, C., Misuraca, G., Mortati, M., Rizzo, F. and Timan, T., AI Watch - Artificial Intelligence for the public sector - Report of the "1st Peer Learning Workshop on the use and impact of AI in public services", Brussels 11-12 February 2020, Publications Office of the European Union, Luxembourg, 2020, p.10, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120315/jrc120315_jrc120315_ai-watch_public-sector_ws_report.pdf.

[38] European Parliament resolution of 9 June 2016 for an open, efficient and independent European Union administration (2016/2610(RSP)) OJ C 86, 6.3.2018, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016IP0279&from=EN.

[39] Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights, Statement on Visit to the United Kingdom, London, 16 November 2018, p. 7, https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881&LangID=E.

[40] J. Redden, Democratic governance in an age of datafication: Lessons from mapping government discourses and practices p. Big Data & Society, July–December 2018, p. 34.

[41] I. Lindgrena, C. Madsen, S. Hofmann, U. Melin, Close encounters of the digital kind: A research agenda for the digitalization of public services, Government Information Quarterly Volume 36, Issue 3, July 2019, p.432

[42] Leslie, D. (2019), Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector, The Alan Turing Institute, p. 23-24. https://www.turing.ac.uk/research/publications/understanding-artificial-intelligence-ethics-and-safety

mechanisms is also reinforced by the opaqueness of the algorithmic models and outputs, with individuals deprived of critical information to understand and challenge this decision-making.

III.   Features, trends and challenges of regulating AI in the public sector

At the European level, there are several initiatives aiming to promote and facilitate timely and coordinated AI regulation, including AI uses from and within the public sector.[43] In May 2020, the European Parliament, the Legal Affairs Committee discussed three draft reports on artificial intelligence: the draft report on AI civil liability,[44] the draft report on AI ethical framework[45], and the draft report on intellectual property rights for the development of artificial intelligence technologies.[46] Similar developments are observed at the international level in order to intervene promptly and adequately to regulate the AI uses and effects since they do not meet geographical barriers.[47] On 21 April 2021, the European Commission set out a proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain legislative acts. The Proposal is underpinned by the specific objectives of: ensuring AI systems placed on the Unions market and used are safe and respect existing law on fundamental rights and Union values; ensuring legal certainty to facilitate investment and innovation in AI; enhancing governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems; and facilitating the development of a single market for lawful, safe and trustworthy AI applications and preventing market fragmentation.[48]

Governments aim to create comprehensive frameworks and policies in relation to digital regulation to address the deployment of AI solutions across the central government and wider public sector departments. Whether a pre-emptive or reactive measure to control and regulate the development and application of AI in the public sector, these initiatives range from institutionalised binding frameworks and measures to soft-law and non-binding guidance. For example, in 2018, EU Member States signed the Declaration of Cooperation on Artificial Intelligence committing to make AI available

---

[43] See indicatively European Parliament, Civil Law Rules on Robotics, European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html, High-Level Expert Group on AI, Ethics Guidelines for Trustworthy AI, European Commission, 2018, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, and European Parliament, Legislative Train Schedule: Connected Digital Single Market, 2018, https://www.europarl.europa.eu/legislative-train/theme-connected-digital-single-market/file-white-paper-artificial-intelligence-and-follow-up/12-2018.

[44] European Parliament, Draft report on AI civil liability (rapporteur Axel Voss, EPP, Germany), 2020/2014(INL),

[45] European Parliament, Draft report on AI ethical framework (rapporteur Ibán García del Blanco, S&D, Spain) 2020/2012(INL). SIENNA and SHERPA jointly provided feedback to the Rapporteurs of this report on 22 May 2020.

[46] European Parliament, Draft report on intellectual property rights for the development of artificial intelligence technologies (rapporteur Stéphane Séjourné, Renew, France), 2020/2015(INI).

[47] See indicatively Erdélyi, Olivia J., and Judy Goldsmith, "Regulating artificial intelligence: Proposal for a global solution." Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, ACM, 2018. https://par.nsf.gov/servlets/purl/10066933 and https://www.coe.int/web/artificial-intelligence/cahai.

[48] Proposal for a Regulation of The European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts COM/2021/206 Final, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

and beneficial to public administrations,[49] whereas ten governments signed the Declaration on Artificial Intelligence in the Nordic-Baltic Region.[50] According to research conducted by the Observatory of Public Sector Innovation Observatory in 2019, "out of 50 countries (including the European Union) with complete or forthcoming national AI strategies, 36 have either strategies in place for public sector transformation through AI or a dedicated focus embedded within a broader strategy".[51]

For instance, Sweden adopted the National Approach to Artificial Intelligence in June 2018 and established a new Ministry of Infrastructure in January 2019 to oversee the national digitalisation.[52] France adopted the National Artificial Intelligence Strategy in 2018, and one year later the public administration was invited to express its interest to experiment with artificial intelligence in public services.[53] Similar developments were noted in Spain, with strategic policies and frameworks to support the digital government and ICT initiatives, including a National Artificial Intelligence Strategy to boost public and private synergies and coordinate national policies.[54] Among the most AI-inspired States, the UK aims to boost digital governance. It assigned to the Government Digital Service, in liaison with the Office of Artificial Intelligence, the task of examining how AI can be best operationalised in the public sector for enhanced public sector productivity.[55] In a similar vein, Germany is using case-based-reasoning algorithms for alphanumerical searches and has designed a comprehensive strategy to implement AI across the public sector, including uses in law enforcement, internal and external security.[56] At the global level, the OECD AI Policy Observatory[57] was set up in February 2020 and provides an inclusive platform with international data, resources and insights into AI to support public policy on AI. The United Nations also launched the E-Government Survey 2020 in July 2020, which assesses the e-government development status of all United Nations Member States.[58]

Despite the divergences among the national, regional and international strategies and frameworks, there are shared features, patterns and trends, which could be considered the first-generation regulatory indications and approaches. First, stakeholders acknowledge that there is a need to intervene and address the challenges and promising benefits of AI to facilitate the public sector digitalisation, mainly in the areas of health care, technology, agriculture, and manufacturing. Most

---

[49] EU Declaration on Cooperation on Artificial Intelligence, 2018, https://ec.europa.eu/jrc/communities/sites/jrccties/files/2018aideclarationatdigitaldaydocxpdf.pdf.

[50] AI in the Nordic-Baltic region, 2018, https://www.norden.org/en/declaration/ai-nordic-baltic-region.

[51] Jamie Berryhill & Kévin Kok Heang & Rob Clogher & Keegan McBride, 2019, "Hello, World: Artificial intelligence and its use in the public sector," OECD Working Papers on Public Governance 36, OECD Publishing. p. 73, https://www.oecd-ilibrary.org/governance/hello-world_726fd39d-en.

[52] European Commission, 'Digital Government Factsheet 2019: Sweden', 2019, https://joinup.ec.europa.eu/sites/default/files/inline-files/Digital_Government_Factsheets_Sweden_2019.pdf.

[53] European Commission, 'Digital Government Factsheet 2019: France', 2019 https://joinup.ec.europa.eu/sites/default/files/inline-files/Digital_Government_Factsheets_France_2019.pdf.

[54] European Commission, 'Digital Government Factsheet 2019: Spain,' 2019 https://joinup.ec.europa.eu/sites/default/files/inline-files/Digital_Government_Factsheets_Spain_2019_1.pdf.

[55] European Commission, 'Digital Government Factsheet 2019: The United Kingdom,' 2019 https://joinup.ec.europa.eu/sites/default/files/inline-files/Digital_Government_Factsheets_UK_2019.pdf.

[56] The Federal Government. Artificial Intelligence Strategy, November 2018. https://ec.europa.eu/knowledge4policy/ai-watch/germany-ai-strategy-report_en.

[57] OECD AI Policy Observatory <https://oecd.ai/>.

[58] UN E-Government Survey 2020 available at https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2020

national AI strategies balance competing interests and values, such as data innovation and respect of human rights and aim to "modernise public administration, public services and research centres".[59] Second, although national strategies tend to be silent in relation to specific regulatory measures, AI strategies and policies are often underpinned by planned legislative reforms or existing legal acts so that AI does not deregulate the public sector. Indeed, the risk that governors and governing democratic rules and procedures may be substituted by the object of regulation, i.e., AI, should be addressed without undue delay.

Third, there is a general consensus that AI should serve the public mission (and protect the public interest[60]) and not *vice versa*. To this end, these policies often refer to human rights and ethical commitments, including transparency, fairness and non-discrimination.[61] Despite the great value of these commitments, the mere reference to broad and hard to define frameworks and principles may be void of meaning and impact. This risk is more intense in national strategies that are intended to pave the way for AI revolution, but they do not consider the balancing between AI and human rights or weigh conflicting values in the context of AI, such as transparency and privacy against the need for a digital and efficient public sector. At the European level, it was found that "emerging national strategies on AI in the public sector in Europe seem to privilege a 'sermon approach' over sticks and carrots."[62] Indeed, these strategies focus more on soft policy instruments and options, such as awareness-raising, rather than actual implementation solutions, including funding and procurement.

Fourth, it is worth noting that national strategies are established to support digital transformation and foster innovation. In this context, there is not always clarity on whether the regulatory reform will refer to self-regulation, co-regulation, or full regulation, whether this will be sectoral or overarching across the public sector, or whether any new law will be rule-based or principle-based. In the above context, national AI policies are often considered high-level and ill-designed for the real world. Indeed, following the review of 34 national AI policies, it was found "that governments are failing to plan for operational investments, continue to be far more aspirational than practical in their planning, and are failing to consider funding realities."[63] Fifth, national strategies and frameworks aim to boost synergies between the public and private sector and establish innovation hubs and labs.[64] For example, in June 2020, the UK government signed a memorandum of understanding with Google Cloud to promote the adoption of its technologies and services in the UK public sector, including central and regional

---

[59] Craglia M., Annoni A., Benczur P., Bertoldi P., Delipetrev P., De Prato G., Feijoo C., Fernandez Macias E., Gomez E., Iglesias M., Junklewitz H, López Cobo M., Martens B., Nascimento S., Nativi S., Polvora A., Sanchez I., Tolan S., Tuomi I., Vesnic L., Artificial Intelligence: A European Perspective, 2018, p. 108. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/ai-flagship-report-online.pdf.

[60] See, e.g., the Proposal for a Regulation of The European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts COM/2021/206 Final, Recital 5.

[61] Jamie Berryhill & Kévin Kok Heang & Rob Clogher & Keegan McBride, 2019, "Hello, World: Artificial intelligence and its use in the public sector," OECD Working Papers on Public Governance 36, OECD Publishing. https://www.oecd-ilibrary.org/governance/hello-world_726fd39d-en.

[62] P. Misuraca, G., and van Noordt, C.., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, p. 67, https://joinup.ec.europa.eu/sites/default/files/document/2020-07/jrc120399_Misuraca-AI-Watch_Public-Services_30062020_DEF_0.pdf.

[63] S. Fatima, K. Desouza, G. Dawson, National strategic artificial intelligence plans: A mullti-dimensional analysis, Economic Analysis and Policy Volume 67, September 2020, p. 192.

[64] J. Redden, Democratic governance in an age of datafication: Lessons from mapping government discourses and practices, Big Data & Society 2018, p. 1-13.

services and charities.[65] Nonetheless, in the absence of restrictive provisions or robust regulatory frameworks about the AI partnerships between the public and private sector, there is a risk that AI public tools and services are outsourced to companies under opaque regulatory leeway, obscure legal conditions and lack of public scrutiny. In addition, in relation to these public-private bridges, it has been observed that they could result in ethical concerns about the data uses by private companies and the technological deskilling of public sector employees.[66] Another risk regarding the increased and unregulated AI cooperation between the public and private sector relates to obscuring the AI chain. Indeed, the lack of clarity about the number and type of private actors involved in AI public uses and the underpinning relationship between companies and the government could inhibit individuals from holding organisations responsible for AI damage.

Therefore, it is imperative that a fit-for-purpose framework is designed, which builds on our historic policy-making and regulatory experience with emerging technologies, while being agile and appropriate to address the AI challenges and enable the AI public sector transformation with controls and safeguards.

IV.   Framework for regulating AI in the public sector: A proposal of critical elements

AI's power lies in the promising transition for States to sustainable, subject-centric, transparent, responsible and responsive performers of their public duties and missions. In a modern adaptation of *The Castle of Kafka*, the application of transformative technologies to the public sector would amount to the victory of transparency, efficiency and inter-connectivity over bureaucracy. On the other side of the AI continuum, there is a lurking risk that AI will disrupt this public service upgrade resulting in the construction of AI-heavy, bureaucratic, inaccessible, opaque public machines. There is an intense discussion about the integration of AI into human lives, public processes and procedures and there is a need for multidisciplinary cooperation across legal, policy, public engagement and ethics fields.[67]

From the laboratory environment and the antagonistic private realm, AI is establishing its presence within public governance. Whereas AI applications are becoming more visible, omnipresent, affordable, and accessible in the public sphere and individuals are interacting with them, AI is also becoming more and more obscure, opaque and over-dominant. At the same time, AI is deployed both as an emergency measure in dealing with unanticipated threats and risks against societal welfare, such as in public health crises, and a routine application across the public sector or branches of it. For instance, the Covid-19 pandemic accelerated the deployment of AI across the public sector under short deadlines and significant pressures to 'normalise' the fast-track digitalisation of States. The focus has been on efficiency, especially in the healthcare sector[68], rather than on thoroughly planned and

---

[65] Google Cloud lays flag in UK public sector with government MoU, 3 June 2020, https://www.verdict.co.uk/google-cloud-uk-government-mou-public-sector/.

[66] Centre for Public Impact, Destination unknown: Exploring the impact of Artificial Intelligence on Government, 2017, p. 42. https://resources.centreforpublicimpact.org/production/2017/09/Destination-Unknown-AI-and-government.pdf.

[67] See indicatively Oxford Insights, Government Artificial Intelligence Readiness Index, 2019 https://ai4d.ai/wp-content/uploads/2019/05/ai-gov-readiness-report_v08.pdf and Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector, The Alan Turing Institute, https://www.turing.ac.uk/research/publications/understanding-artificial-intelligence-ethics-and-safety.

[68] Sun, Tara Qian, and Rony Medaglia. "Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare." Government Information Quarterly 36.2 (2019), p. 368-383.

assessed AI applications, often neglecting or underestimating the AI impact on the public and governance structures and long-term effects.[69]

Therefore, there is a need for verifying genuine AI beneficence; AI is missioned to support social good and is designed under safeguards. We need to guard against AI uses that generate abuse and misuse of the public powers.[70] These AI systems should be distinguished from pure "AI for good", where the impact of AI is measured against the intended direct outcome. For instance, with regard to the AI-driven credit-rating systems, the mere enablement of processing applications does not suffice to qualify this safe and beneficent AI. On the contrary, AI should be a regulated object and activity from its inception to its final deployment and interaction with the end-user. Regulation should also cover the second-generation effects of AI and the impact on individuals and social groups other than the planned application field and scope of AI systems.

This section presents a framework for regulating AI in the public sector bringing together what we have identified as critical elements to ensure that the deployment and uses of AI in the public sector are beneficent and socially responsible. The framework covers three core aspects (i) Legal (ii) Ethical and (iii) Societal.

Legal relates to legal compliance and includes measures required to comply with the law. The reference to law indicates adhering to the applicable legal requirements and standards, as they are laid down by international or regional law. In line with the above, the legal territory of AI includes technology-specific legislation, but it is not restricted to commonly discussed legal fields, such as data protection and intellectual property. This territory is not delineated and regulators should constantly assess how AI applications transform the way States and individuals understand access to public services, the performance of public duties and their impact on fundamental rights and freedoms.

Ethical includes elements that address ethical issues and help address these early on, such as ethics by design, ethical impact assessments and use of ethical standards. Ethical components do not substitute the legal elements and requirements, but they support the reflection and inclusion of ethical considerations and the public's perceptions into public policy-making and decision-making.

Societal covers elements critical to minimising societal harm and maximising public benefit. This is rather challenging as both benefits and harms are hard to quantify when it comes to the AI risks, including, for example, the highly discussed risks of AI discrimination. Even where governments are willing to acknowledge the AI risks, regardless of how remote and marginal they may often seem, there is not yet an 'algorithm' to accurately weigh and balance public harms and public benefits. Therefore, there is a need for societal drivers and guiding questions that support this assessment, without sticking to the traditional concepts of harm and benefit, but pre-emptively addressing these questions within the context of AI applications.

This framework will help advance the governance of the use of AI in the public sector and provide insights into the critical elements that need to be further developed if the full potential of AI use in

---

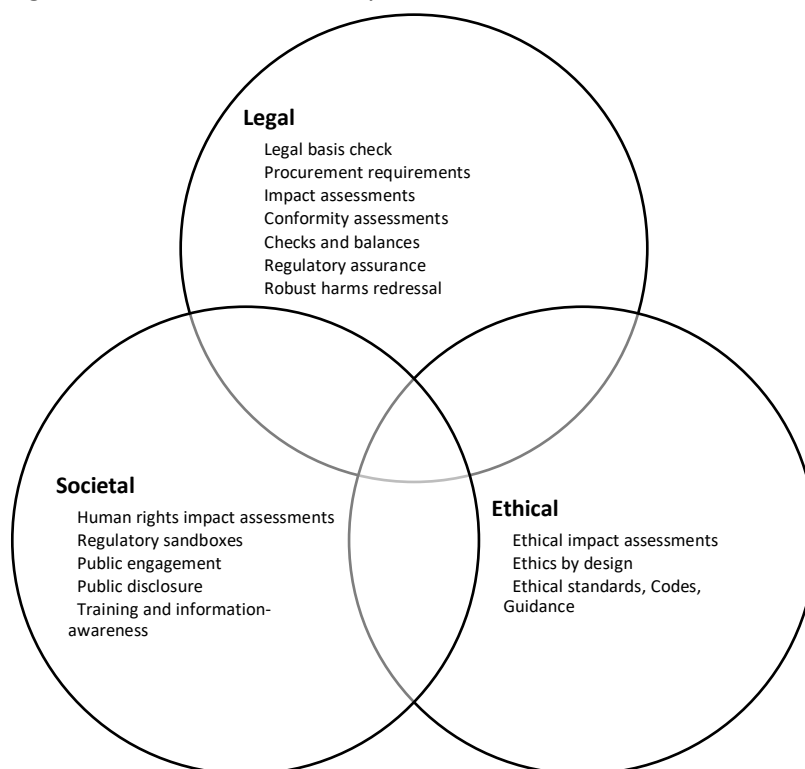[69] Misuraca, G., and van Noordt, C., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, p. 39, 50.

[70] Panel for the Future of Science and Technology, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, June 2020 p. 20, https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2020)641530.

the public sector is to be realised in a manner that meets societal expectations and is in line with fundamental values and freedoms.

The suggested core aspects of this framework should complement, underpin and operationalise the national AI strategies and measures to ensure that AI risks and harms are regulated in a responsible manner, especially in the state-citizen relationship that now displays even more endemic power imbalances. AI should not be exempt from existing governance frameworks in any manner or form, particularly given the high stakes and potential for impact on human life and society from every direction. The Figure below illustrates the key elements of the Framework.

Legal
Legal basis check
Procurement requirements
Impact assessments
Conformity assessments
Checks and balances
Regulatory assurance
Robust harms redressal

Societal
Human rights impact assessments
Regulatory sandboxes
Public engagement
Public disclosure
Training and information-awareness

Ethical
Ethical impact assessments
Ethics by design
Ethical standards, Codes, Guidance

Fig 1: Framework for regulating AI in the public sector

Many of the outlined elements are not exclusive to one or the other and may overlap in functionality, scope and purpose. The above is not a comprehensive or one-off statement of what should be; the elements have the potential for expansion based on AI, policy and regulatory advances. These elements were drawn from our research into the regulatory options for AI[71] - however, these have been outlined and reiterated in many other places as illustrated in the table below[72]. [In the table: A = academic/research paper; CS = Civil society report; E= Ethical; G= Guideline; L = Legal; P = Policy document; M = Media article; R = Regulation/Legislation/Similar measure; RP= Research Project; S= Societal]

Element	Documented/expressed in, e.g.,

---

[71] R Rodrigues, A Panagiotopoulos, B Lundgren, S Laulhé Shaelou, A Grant (2020): D3.3 Report on regulatory options. SHERPA. <https://doi.org/10.21253/DMU.11618211.v3>

[72] Note these are non-exhaustive examples of where the elements have been featured.

Explicit legal basis (L)	P, [73]R[74]
Procurement requirements (L)	P[75], R[76]
Impact assessment (L)	A[77], P[78], R[79]
Conformity assessments (L/S)	P[80], RP[81]
Checks and balances (L)	A[82], P[83], R[84]
Regulatory assurance (L)	A[85], P[86]

---

[73] E.g., UN General Assembly, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, 9 April 2013, https://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf

[74] E.g., European Parliament resolution on autonomous weapon systems (2018/2752(RSP); European Parliament, European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2013(INL)), P8_TA(2017)0051; "DEEP FAKES Accountability Act", (H.R. 3230) 116th Cong. (2019) Proposer/Sponsor: Rep. Yvette D. Clarke, [D-NY-9] (Introduced on 12 June 2019), https://www.congress.gov/bill/116th-congress/house-bill/3230/text.

[75] High-Level Expert Group on AI (AI HLEG), Policy and investment recommendations for trustworthy Artificial Intelligence, 26 June 2019, https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

[76] E.g., Government of Canada, Directive on Automated Decision Making, effective 1 April 2019, https://www.tbs-sct.gc.ca/pol/doc- eng.aspx?id=32592.

[77] Kaminski, Margot E. and Malgieri, Gianclaudio, "Algorithmic Impact Assessments under the GDPR: Producing Multi-layered Explanations", U of Colorado Law Legal Studies Research Paper No. 19-28, http://dx.doi.org/10.2139/ssrn.3456224.

[78] High-Level Expert Group on AI (AI HLEG), Policy and investment recommendations for trustworthy Artificial Intelligence, 26 June 2019, https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

[79] E.g., Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119, 4.5.2016, p. 1-88.

[80] European Commission, White paper On Artificial Intelligence - A European approach to excellence and trust, COM (2020) 65, February 2020.

[81] Siemaszko K, Rodrigues R, Slokenberga S, D5.6, Recommendations for the enhancement of the existing legal frameworks for genomics, human enhancement, and AI and Robotics. SIENNA, 2020. https://www.sienna-project.eu/digitalAssets/894/c_894270-l_1-k_sienna_d5.6_recommendations-for-the-enhancement-of-the-existing-legal-frameworks-for-genomics--human-enhancement--and-ai-and-robotics_www.pdf.

[82] Colin Gavaghan, Alistair Knott, James Maclaurin, John Zerilli, Joy Liddicoat, Government use of artificial Intelligence in New Zealand, Final Report on Phase 1 of the New Zealand Law Foundation's Artificial Intelligence and Law in New Zealand Project, 2019, https://www.cs.otago.ac.nz/research/ai/AI-Law/NZLF%20report.pdf, Erdélyi, Olivia J., and Judy Goldsmith, "Regulating artificial intelligence: Proposal for a global solution." Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, ACM, 2018, https://par.nsf.gov/servlets/purl/10066933.

[83]Session of the UN General Assembly, 2013.

[84] European Parliament Resolution 2018; European Commission For The Efficiency Of Justice (CEPEJ), Charter on the use of Artificial Intelligence in judicial systems and their environment. Adopted at the 31st plenary meeting of the CEPEJ, Strasbourg, 3-4 December 2018, https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

[85] Tutt, Andrew. "An FDA for algorithms." Admin. L. Rev. 69 (2017): 83.

[86] Committee on Standards in Public Life, Artificial Intelligence and Public Standards, 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF.

Ethical impact assessments (E)	A[87], RP[88],
Ethics by design (E)	A[89], P[90], RP[91]
Ethical standards, Codes, Guidance (E)	P[92]
Human rights impact assessments (S)	P[93]
Regulatory sandboxes (S/L)	CS,[94] P[95], RP[96]
Public engagement (S)	RP[97]
Public disclosure (S)	A[98], P[99],
Robust harms redressal (S/L)	A[100], P,[101] RP,

Each of the individual elements of the framework has been subject to lengthy and rigorous policy, academic and legal analysis, and so we do not seek to duplicate this here. Rather, we next move to break out these elements into evaluative and reflexive assessment questions that will be useful to **policy-makers (and researchers evaluating AI policy in the public sector)** particularly at the national

---

[87] Mantelero, Alessandro. "AI and Big Data: A blueprint for a human rights, social and ethical impact assessment." Computer Law & Security Review 34.4 (2018): 754-772; Wright, D. 2011. A framework for the ethical impact assessment of information technology. 13 Ethics Inf. Technol. 199, p. 201–202.

[88] Siemaszko K, Rodrigues R, Slokenberga S, D5.6, Recommendations for the enhancement of the existing legal frameworks for genomics, human enhancement, and AI and Robotics. SIENNA, 2020.

[89] Iphofen, Ron, and Mihalis Kritikos. "Regulating artificial intelligence and robotics: ethics by design in a digital society." Contemporary Social Science (2019), p. 1-15.

[90] Committee on Standards in Public Life, 2020.

[91] Siemaszko K, Rodrigues R, Slokenberga S, D5.6, Recommendations for the enhancement of the existing legal frameworks for genomics, human enhancement, and AI and Robotics. SIENNA, 2020.

[92] European Parliament, Draft report on AI ethical framework (rapporteur Ibán García del Blanco, S&D, Spain) 2020/2012(INL), 21.4.2020, https://www.europarl.europa.eu/doceo/document/JURI-PR-650508_EN.pdf

[93] The Council of Europe Commissioner for Human Rights, "Unboxing Artificial Intelligence: 10 steps to protect Human Rights", 2019, p. 6, https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect- human-rights-reco/1680946e64.

[94] Access Now (2018) Mapping Regulatory Proposals for Artificial Intelligence in Europe, https://www.accessnow.org/mapping-artificial-intelligence-strategies-in-europe/

[95] European Commission, Annex to the Coordinated Plan on Artificial Intelligence (COM(2018) 795) 7 December 2018, https://ec.europa.eu/digital-single- market/en/news/coordinated-plan-artificial-intelligence; https://iapp.org/news/a/norwegian-dpa-creating-regulatory-sandbox-for-ai/.

[96] R Rodrigues, A Panagiotopoulos, B Lundgren, S Laulhé Shaelou, A Grant (2020): D3.3 Report on regulatory options. SHERPA, https://doi.org/10.21253/DMU.11618211.v3.

[97] Whittlestone, J. Nyrup, R. Alexandrova, A.Dihal, K. Cave, S. (2019) Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. London: Nuffield Foundation.

[98] Colin Gavaghan, Alistair Knott, James Maclaurin, John Zerilli, Joy Liddicoat, Government use of artificial Intelligence in New Zealand, Final Report on Phase 1 of the New Zealand Law Foundation's Artificial Intelligence and Law in New Zealand Project, 2019, https://www.cs.otago.ac.nz/research/ai/AI-Law/NZLF%20report.pdf.

[99] European Commission For The Efficiency Of Justice (CEPEJ), Charter on the use of Artificial Intelligence in judicial systems and their environment, 2018.

[100] Quintarelli, S, We need "redress by design" for AI systems, Quinta's weblog, 8 April 2019. https://blog.quintarelli.it/2019/04/we-need-redress-by- design-for-ai-systems.html.

[101] See European Commission For The Efficiency Of Justice (CEPEJ), Charter on the use of Artificial Intelligence in judicial systems and their environment and High-Level Expert Group on Artificial Intelligence (AI HLEG), Ethics Guidelines for Trustworthy AI, 2019.

level to evaluate policy effectiveness and any relevant gaps and challenges. These assessment questions aim to guide regulators, policy-makers and researchers through the key components and concepts of protective AI frameworks, without dictating a regulatory panacea for AI. Their value lies exactly in drawing the regulatory and research community's attention to the critical building blocks of AI regulation and aims to materialise this in a practical and accountable manner.

This is not a 'blacklist' or closed catalogue and the assessment questions should be adapted to the national, political and legal idiosyncrasies of States, without thwarting their very own purpose, i.e., to control and regulate AI uses in a proportionate and accountable manner. This is particularly important given the commitments of governments and the raised expectations "for governments to play a more prevalent role in the digital society and to ensure that the potential of technology is harnessed, while negative effects are controlled and possibly avoided" particularly to avoid harm to democracy and trust in governance systems.[102]

The assessment questions have been drawn from the range of literature studied during the course of this research, policy and academic at different levels (international, regional - particularly EU level- and national). They were also particularly inspired by the work of the authors in the SHERPA and SIENNA Horizon 2020 project – specifically the SHERPA study of regulatory options,[103] and the SIENNA recommendations for enhancement of legal frameworks for AI and robotics.[104]

ASSESSMENT QUESTIONS

Legal

    a. **Explicit legal basis**
        i. Is there an explicit legal basis for the development, training and trialling of the AI system/product/service?
        ii. Do the AI deployment and use in the public sector have an explicit legal basis so that it does not undermine public standards and the rule of law?
        iii. Are there provisions for checks to be carried out to ensure that the deployment and use of AI comply with relevant laws and regulations prior to such deployment/use and to ensure the deployment is appropriate,[105] necessary, proportionate and lawful?
        iv. Is there guidance on harmful AI deployment and uses in the public sector?
        v. Are there other legal obligations and frameworks in relation to the implications of AI applications in the public sector that are not covered by the primary permissive legal basis?

---

[102] Kuziemski, Maciej, and Gianluca Misuraca. "AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings." Telecommunications Policy (2020): Jul; 44(6): 101976. doi: 10.1016/j.telpol.2020.101976.

[103] R Rodrigues, A Panagiotopoulos, B Lundgren, S Laulhé Shaelou, A Grant (2020): D3.3 Report on regulatory options, SHERPA, https://doi.org/10.21253/DMU.11618211.v3.

[104] Siemaszko K, Rodrigues R, Slokenberga S, D5.6, Recommendations for the enhancement of the existing legal frameworks for genomics, human enhancement, and AI and Robotics. SIENNA, 2020.

[105] See UNESCO Ad Hoc Expert Group (AHEG) for the preparation of a draft text of a recommendation on the ethics of artificial intelligence SHS/BIO/AHEG-AI/2020/4 REV.2 Paris, 7 September 2020, https://unesdoc.unesco.org/ark:/48223/pf0000373434.

b. Procurement requirements

 i. Are there national/local/departmental Directives/regulations/policies guidelines and set requirements for the procurement of AI systems, products and services?

 ii. Are these national/local/departmental procurement guidelines in line with principles of good administration, accountability and transparency?

 iii. Are there provisions for public accountability and scrutiny of AI applications?

 iv. Are there obligations/provisions to make the details of the suppliers of AI tools and the relevant agreements public?

 v. Does the AI procurement framework regulate the entire lifecycle of the procurement and contract, including any further uses of the training data sets and AI outputs by the supplier after the termination or expiry of the supplier agreement?

 vi. Are public sector organisations required to put appropriate measures in place to verify and select reliable third parties and further implement appropriate safeguards and controls in the supply chain and procurement process?

 vii. Are there provisions for AI risk assessments to demonstrate areas of risks and vulnerabilities when involving private entities in using public resources (e.g., information assets and personal data) for AI applications?

 viii. Are suppliers adhering to codes of practice and recognised standards and providing AI and data protection training to their employees?

 ix. Are public sector organisations preparing standard selection questionnaires and due diligence questions to verify that the selected supplier continues to conform to the relevant requirements?

 x. Are appropriate data governance mechanisms being designed in advance to enable safe data sharing through dedicated secure platforms and mechanisms?

 xi. Are there obligations for auditing suppliers and their partners during and following the termination or expiry of the AI partnership to ensure compliance with applicable law and the contractual terms?

c. Impact assessments: structured process to identify and evaluate the likely positive and negative impacts of a proposed project or development.

 i. Are there policy and/or legal incentives/requirements supporting the use of legally required impact assessments (e.g., data protection impact assessments) to assess the risks and benefits of AI development, deployment and use – taking into account ethical, environmental, societal, human rights risks and impacts?

 ii. Are the requirements well-timed, e.g., at the start of the procurement process, with provision for re-visiting at key decision points?

d. Conformity assessments

 i. Are there requirements in place for conformity assessments (mandatory or voluntary), including procedures for AI testing, inspection or certification, such as checks of the algorithms and of the data sets used?

 ii. Is there a provision/process in place to evaluate the need for such measures and actions to be taken, if required?

    iii.   Are there agreed means and procedures for: a) enforcing compliance; and b) remedying non-conformity should these assessments raise concerns?

e. **Checks and balances**: safeguards and measures under existing legal frameworks to prevent power imbalances, foster power-sharing, cooperation and oversight.

    i.   Is there a legislative framework for independent and effective oversight over the human rights compliance of the development, deployment and use of AI systems by public authorities and private entities, which includes mechanisms that consist of a combination of administrative, judicial, quasi-judicial and/or parliamentary oversight bodies effectively cooperating with each other?[106]

    ii.   Are there provisions for checks and balances to maintain the rule of law and public trust in the government deployment and use of AI in the public sector?

    iii.   Is there a provision for a Register of algorithms[107] used in government?

    iv.   Is there provision for judicial review of the AI-made decisions to establish legal clarity, where required?

    v.   Are there provisions for traceability mechanisms of the data used for the training and operation of AI systems?

    vi.   Are there provisions for traceability mechanisms of the uses and outputs of AI systems?

    vii.   Are there provisions for obligatory use of synthetic data in the training stage of the algorithmic model where personal data is not necessary?

    viii.   Is there a regulated and monitored ecosystem to enable public sector organisations to cooperate and rely on pre-approved technical standards, architectures and functionalities, anonymised sets of data, registers of approved vendors, and a list of cyber security requirements?

    ix.   Are there robust record-keeping requirements for data sets used to train and test the AI systems etc.?

    x.   Are there risk management provisions for AI vigilance systems (e.g., algorithmic vigilance similar to pharmacovigilance mechanisms to contain unforeseen risks and outcomes)?[108]

f. **Regulatory assurance**: includes regulatory quality monitoring activities.

    i.   Are there provisions to evaluate the gaps in the regulatory landscape on a regular basis (given the changing policy and technological landscape) and advise regulators and government on the issues associated with AI?

    ii.   Is there a national agreement/overarching framework to provide regulatory assurance?

---

[106] The Council of Europe Commissioner for Human Rights, "Unboxing Artificial Intelligence: 10 steps to protect Human Rights", 2019, https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect- human-rights-reco/1680946e64.

[107] See e.g., Gavaghan, Colin, Alistair Knott, James Maclaurin, John Zerilli, Joy Liddicoat, Government use of artificial intelligence in New Zealand, Final Report on Phase 1 of the New Zealand Law Foundation's Artificial Intelligence and Law in New Zealand Project, 2019, https://www.lawfoundation.org.nz/wp-content/uploads/2019/05/2016_ILP_10_AILNZ-Report-released-27.5.2019.pdf.

[108] Consultative Committee Of The Convention For The Protection Of Individuals With Regard To Automatic Processing Of Personal Data (Convention 108), Report on Artificial Intelligence Artificial Intelligence and Data Protection: Challenges and Possible Remedies, 2019, p. 17, https://rm.coe.int/artificial-intelligence-and-data-protection-challenges-and-possible-re/168091f8a6.

      iii.   Are there provisions via a dedicated body to identify gaps in the regulatory landscape and advise the government and regulators on AI for sector-based practices?

      iv.   Is there an independent regulatory authority to report on the uses of AI in the public sector and provide safeguards for regulatory intervention and judicial remedies on an *ad hoc* or regular basis?

g. **Robust harms redressal**:[109] The mere existence of legal tools and redress mechanisms does not guarantee effective harm redressal. Such mechanisms should be underpinned by societal conditions that enable seamless access to these tools by individuals.[110]

      i.   Do providers of public services (public and private) inform individuals of their rights and method of appeal against automated and AI-assisted decisions?

      ii.   Are public bodies enabling individuals to challenge decisions and seek redress using procedures that are independent and transparent?

      iii.   Is there a meaningful process of redress to enable public bodies to find out what failed and how that failure could be rectified?

Ethical

a. **Ethical impact assessments**: a process of judging the ethical impacts of research and innovation activities, outcomes and technologies that incorporates both means for a contextual identification and evaluation of these ethical impacts and development of a set of guidelines or recommendations for remedial actions aiming at mitigating ethical risks and enhancing ethical benefits, typically in consultation with stakeholders.[111]

      i.   Are there provisions to support ethical impact assessments for AI policies, products, services, projects or programmes that facilitate the consideration and mitigation of ethical issues?

      ii.   Are there provisions for independent AI Ethics Officers to oversee ethical impact assessments and/or give support to this process?[112]

b. **Ethics by design**: provisions for the systematic inclusion of ethical guidelines, recommendations and considerations into the design and development processes.

      i.   Are public sector bodies or the outsourced private entities deploying/using AI required to adopt ethics by design?

---

[109] Questions based on recommendations from the Committee on Standards in Public Life, Artificial Intelligence and Public Standards, 2020, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF.

[110] This element is of dual nature. It also fits under the Societal elements. The sub-questions of this element should be approached and addressed keeping this in mind.

[111] As defined in CEN Workshop Agreement, Ethics assessment for research and innovation - Part 2: Ethical impact assessment framework. CWA 17145-2:2017, https://standards.cen.eu/dyn/www/f?p=204:110:0::::FSP_PROJECT:65088&cs=15985768788487A731AC76A9942EDA697.

[112] UNESCO Ad Hoc Expert Group (AHEG) for the preparation of a draft text of a recommendation on the ethics of artificial intelligence SHS/BIO/AHEG-AI/2020/4 REV.2 Paris, 7 September 2020.

      ii. Are there provisions for embedding ethics by design approach into public sector funded/targeted AI Research and Development?

      iii. Do procurement standards encourage the use of ethics by design?

c. **Ethical standards, codes, guidance**:

      i. Are there established ethical standards, codes and guidance for the development and use of AI in the public sector? If so, are these evaluated and updated regularly to remain fit for purpose?

      ii. Is their use incentivised through legislation mandating/encouraging their use and/or procurement policies?

      iii. Are there good models identified and encouraged for use (from the international or regional levels) e.g., FAST SUM principles, OECD AI principles, ALTAI - The Assessment List on Trustworthy Artificial Intelligence, SHERPA Guidelines for Ethical Development and Use of SIS[113]?

Societal

a. **Human rights impact assessments:** a process to discover, measure and/or map human rights impacts and risks.[114]

      i. Is there a requirement/are public authorities encouraged to carry out human rights impact assessments (HRIAs) on AI systems acquired, developed and/or deployed by them?

      ii. Is there a legal framework to set out the procedure for carrying out HRIAs?

      iii. Does the assessment evaluate the potential impact of the AI system on human rights taking into account the nature, context, scope, and purpose of the system?

      iv. Where a public authority has not yet procured or developed an AI system, is this assessment carried out prior to the acquisition and/or development of that system?

      v. Do the HRIAs include a meaningful external review of AI systems, either by an independent oversight body, e.g., National Human Rights Structures (NHRSs) or an external researcher/auditor with relevant expertise?

b. **Socio-economic impact assessments**: "a structured way of showing a proposal's advantages and disadvantages for society as a whole and for various parties"[115] – they weigh socio-economic costs against socio-economic benefits.

      i. Are socio-economic impact assessments encouraged and/or required?

      ii. In particular, are there provisions to assess and alleviate specific impacts on vulnerable groups (e.g., children, migrants, homeless, those accessing welfare

---

[113] Brey, Philip; Macnish, Kevin; Ryan, Mark (2020): Guidelines for the Ethical Development of AI and Big Data Systems: An Ethics by Design approach, SHERPA, https://doi.org/10.21253/DMU.12301322.v1.

[114] Based on recommendations by The Council of Europe Commissioner for Human Rights, "Unboxing Artificial Intelligence: 10 steps to protect Human Rights", 2019, https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect- human-rights-reco/1680946e64.

[115] Swedish Environmental Protection Agency, Socio-economic impact assessment, http://www.swedishepa.se/Environmental-objectives-and-cooperation/Swedish-environmental-work/Work-areas/Socio-economic-impact-assessment/.

        services) that may experience socio-economic disadvantage from the AI deployment and use*?*

c. **Regulatory sandboxes**: frameworks that allow regulators to carry out "small scale, *live testing of innovations* by private firms in a *controlled environment* (operating under a special exemption, allowance, or other limited, time-bound exception) under the regulator's supervision".[116]

    i. Is good use being made of regulatory sandboxes to determine which AI products, services and systems can be rolled out in the public sector and what regulatory changes are then required?

    ii. Are critical success factors, such as the careful design of parameters, transparency in design, operation and outcomes and good communication and cooperation with stakeholders addressed?

d. **Public consultation and engagement**: engaging the public on AI deployment and use at the appropriate time using principles of effective engagement e.g., transparency, integrity, issue-focused, good design, respect.

    i. Are there good opportunities for collaborative and issue-focused engagement and discussion by the public on the ethical, societal and human rights concerns of AI in the public sector?

    ii. Has a shared public understanding of AI been developed?

    iii. Why types of methods of public consultation/engagement are in use to discuss AI impacts in public sector use? E.g., surveys, forums, public dialogue?

    iv. Do the consultations provide an opportunity for all stakeholders, including state actors, private sector representatives, academia, civil society actors, the media and representatives from marginalised and affected communities to be heard?

    v. Can public objections to AI deployment and use be registered in other ways, e.g., petitions?

e. **Public disclosure**: making information available widely by publishing it in clear and accessible terms.[117] Disclosure about AI systems is critically connected to transparency and accountability and a safeguard of democratic governance.

    i. Is proactive disclosure with regard to AI deployment and use encouraged, e.g., by promoting openness and transparency by design?

    ii. Are there government-issued guidelines on this?

    iii. Are there specific freedom of information restrictions in relation to AI systems that should be considered before the deployment of the AI tool or release of critical information to the public?

    iv. Are the AI features communicated in plain language without legal or technical jargon to ensure that the layperson understands and foresees the AI uses and implications?

f. **Training and information awareness**

---

[116] Jenik, Ivo and Kate Lauer, Regulatory Sandboxes and Financial Inclusion, CGAP Working Paper, October 2017, https://www.cgap.org/sites/default/files/Working-Paper- Regulatory-Sandboxes-Oct-2017.pdf.

[117] See The Council of Europe Commissioner for Human Rights, "Unboxing Artificial Intelligence: 10 steps to protect Human Rights", 2019, https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect- human-rights-reco/1680946e64.

      i.   Is training and educational material provided to public officers to understand the basic components, advantages, risks and impacts of AI?

     ii.   Are there informative tools/modules on AI impacts, publicly available and free of charge for individuals, to support them in establishing their rights against decisions made or facilitated by AI?

The value

Contrary to inflexible provisions, open-ended principles and over-rigid frameworks that fail to address the novelty of AI and control its uses in a regulated environment, we suggest the adoption and further tailoring of these assessment questions. These questions do not purport to dictate or impose a pre-determined regulatory outcome or stance against AI. They reflect existing regulatory requirements, policy recommendations and best practices in AI regulation (and other innovative technologies) and aim to support stakeholders in their decision-making. At the same time, they are not neutral, but they tend to incorporate the state of the art in AI regulation taking into account the obligations arising from democratic governance and, especially, human rights frameworks. The way they are framed allows for regulatory considerations, public scrutiny and self-reflection. It also supports narrative and detail; it is not intended to be used just as a checklist but as a guiding and empowering instrument for AI research, policy-making and regulation. Their value also lies in their flexibility to be adapted and implemented within the current regulatory systems and frameworks. They do not aim to substitute the current legislation and regulatory principles, but – set within the democratic spirit- aim to support stakeholders and the public sector to remain agile and flexible in regulating and addressing the impact of AI. They also pave the way for enhanced principles-based AI regulation in combination with hard legal rules to ensure the timeliness, relevance and appropriateness of AI frameworks.

In particular, the above questions will help policy-makers and interested researchers assess and evaluate the regulatory state of the art; whether core elements have been put in place, the state of affairs, whether they are fit for purpose, what is missing and what further needs to be put in place and/or boosted in terms of regulating AI in the public sector. It will help identify how current regulatory measures are working and any unanticipated effects that need addressing to achieve policy objectives. As these are interrogative questions, they can help authorities to scope the various legal fields and requirements that would apply to AI systems in the public sector. Following this, regulatory loopholes could be addressed with the help of these questions, as these questions may be of guiding value, but they still indicate the potential areas for regulation, monitoring and intervention.

These questions could also support public sector employees in familiarising themselves with core AI responsibilities and requirements without delving into highly doctrinal discussions. The questions can be used for information awareness and training purposes, whereas they also support accountability and transparency in the performance of public duties. In this context, it has been argued that there is a need for a "Big Data due progress", under which public bodies would train and educate employees on AI biases and ensure that there are built-in safeguards for reviewing automated decisions and computer systems and keeping audit trails.[118] Public sector staff may feel more encouraged and included in the process of managing AI in their public remit as they have this toolbox to help them interact with other key stakeholders and understand their key responsibilities, without feeling overwhelmed by legal and AI jargon.

---

[118] Edwards, L. and Veale, M. (2017). Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. Duke Law and Technology Review, p. 16-84.

We also hope that these questions will be customised by the relevant authority or stakeholder. In this context, the open-ended character of these questions enables their tailoring to specific contexts (e.g., policy making, judicial or law enforcement) and the focus on specific elements where required (e.g., transparency when the public and affected individuals have an enhanced right and interest to know about the applied AI tool). As these questions pertain to different stages of AI development and deployment, they also provide metrics and performance indicators about the best measures to implement and when. These questions will be raised most probably by different actors at different intervals and this could facilitate a thorough assessment of the different particles of the AI lifecycle. At the same time, this temporal element also renders these questions tools of cooperation, as different actors in the AI public sector chain will need to cooperate to exchange information and review their practices. This is a salient secondary effect of these questions as they intend to be seen as synergistic and relationship-building questions depending on the context and specific circumstances rather than a standalone checklist to be filled only by a sole stakeholder.

In the long-term, these questions could support robust and evidence-based regulation, preventing opportunistic regulatory interventions and knee-jerk regulatory responses that are short-sighted and do not set balanced and sustainable rules. In our previous research, we assessed different approaches to regulating AI and summarised our key findings as per below:

> "Although there are disagreements, important actors seem to **aim for harmonized rules** (as indicated, e.g., by the OECD guidelines and the major companies behind Partnership on AI). Most proposals **advocate a heavier rather than lighter touch**, but there are clear disagreements. Proposals often **combine risk-based approaches with principle-based regulation** (perhaps because, e.g., safety concerns are often risk-related, while some human rights require a principle-based approach). There is an understanding among industrial proponents that **regulations are needed, but there is disagreement and ambiguity** as to whether this should be self-regulation, co-regulation, or full regulation. Unsurprisingly, the most common worry is that **a heavy-touch will restrict innovation**, while a **light-touch will leave individuals and society exposed to risks to fundamental values or human rights.** The challenge for any regulation is how to promote good AI development and use, and how to minimize the creation of bad AI or misuse of AI-technology and increase its security (reliability and resilience)."[119]

These assessment questions could bridge the above-mentioned gaps and act as preparatory work for transitioning to new models of interactive and grounded policy-making and regulation of AI in the public sector.

When to carry out this assessment

As these questions may be addressed by different stakeholders as preventative or repressive safeguards, we would not suggest a specific timeframe for their implementation. It is important to ensure that all the stages of the AI lifecycle are covered by the questions and that there is a proactive approach to them. These questions should not be triggered to justify or legalise the use of AI retrospectively or use in any way to remove barriers and responsibilities in AI regulation. On the contrary, their application should reinforce current frameworks and not impoverish their provisions

---

[119] R. Rodrigues, A. Panagiotopoulos, B. Lundgren, S. Laulhé Shaelou, A. Grant
(2020): D3.3 Report on regulatory options, SHERPA, p. 27, https://doi.org/10.21253/DMU.11618211.v3.

or challenge their application to the extent that they support the above-mentioned democratic principles and values.

To ensure this effect, the assessment should be timely, well-planned and periodic. For example, the assessment could be carried out every year or every two years so that the policy-maker has relevant and timely information that benefits the regulation of the use of AI in the public sector. The depth of the analysis should reflect the significance of desired changes. This is also contingent on the AI technological developments, the urgency of the regulatory intervention and the findings provided through this assessment. It would be also recommended to set a review period for this assessment or, at least, for the elements where stakeholders have identified loopholes and vulnerabilities.

How

Although these are assessment questions, they aim to support regulatory improvements and provide actionable findings. The least interventionist approach would be for the public sector to enjoy the flexibility to revise their own practices and implement additional safeguards. These questions could identify tailored solutions that fit within the regulatory frameworks (e.g., transparency mechanisms, procurement checks) or that could improve best practices and regulatory requirements (e.g., impact assessments). A more radical effect would include triggering a legislative change in line with the assessment's findings. This will depend on the legal system of the State. The scope and framing of these questions support different solutions and facilitate adaptation to the extent that the purpose of this tailoring is not abusive, but it aims at enhancing - and not jeopardising - protection.

The results of the evaluation carried out should be thrown open for consultation, discussion and debate with experts and stakeholders so that they have a chance to provide their feedback and inputs (this should be actively encouraged given the high-risk implications of the use of AI in the public sector). A minimum of three months is recommended. After the consultation, the evaluation findings should be revised and submitted to Parliament/relevant authorities/bodies for consideration and appropriate follow-up actions – particularly to feed into the policy decision-making cycle.

Another important aspect relates to the need to reserve and allocate time to conduct this assessment. Governments should recognise the value of such assessments and provide for the necessary budget, support and expertise. Otherwise, there is a risk that this assessment could become a time-consuming bureaucratic task or a box-ticking exercise.

These questions should not be triggered to justify or legalise the use of AI retrospectively or used in any way to remove barriers and responsibilities in AI regulation. Their application should reinforce current frameworks and not impoverish their provisions; it could challenge their application to the extent that they support democratic principles and values.

V.   Conclusion

As seen here, the AI-driven digitalisation of the public sector comes with great benefits, but as an innovative technology, poorly regulated in its infancy, it could challenge the way the States understand their role, functions and commitments to the public. It is imperative that a fit-for-purpose framework is designed, which builds on our historic policy-making and regulatory experience with emerging technologies, while being agile and appropriate to address the AI challenges and enable the

AI public sector transformation with controls and safeguards. Frameworks for regulating AI in the public sector remain a work in progress and we hope the research here presents a good model and inspiration. It brings together some critical elements to ensure that the deployment and uses of AI in the public sector do not prevent States from performing their obligations and duties towards society and individuals. The assessment questions presented will help advance the governance of the use of AI in the public sector and provide insights into the critical elements needed to be further developed.

We also hope that the assessment questions will be customised for specific contexts and used by the relevant authority and/or stakeholders especially at the national level where AI policies are peaking or where they are in their infancy. The questions will support robust and evidence-based AI regulation, preventing opportunistic regulatory interventions and knee-jerk regulatory responses.

The Framework, including the assessment questions presented here, need further review, testing and development – we expect users will have/bring their own expectations and requirements also to this assessment. In this sense, the questions should not be seen as a static or fixed mechanism. AI itself is severely volatile, disruptive and expanding in the public sector. This Framework and its questions must mirror this aspect and themselves be agile, reflexive and dynamic and be used in this manner.

Acknowledgements