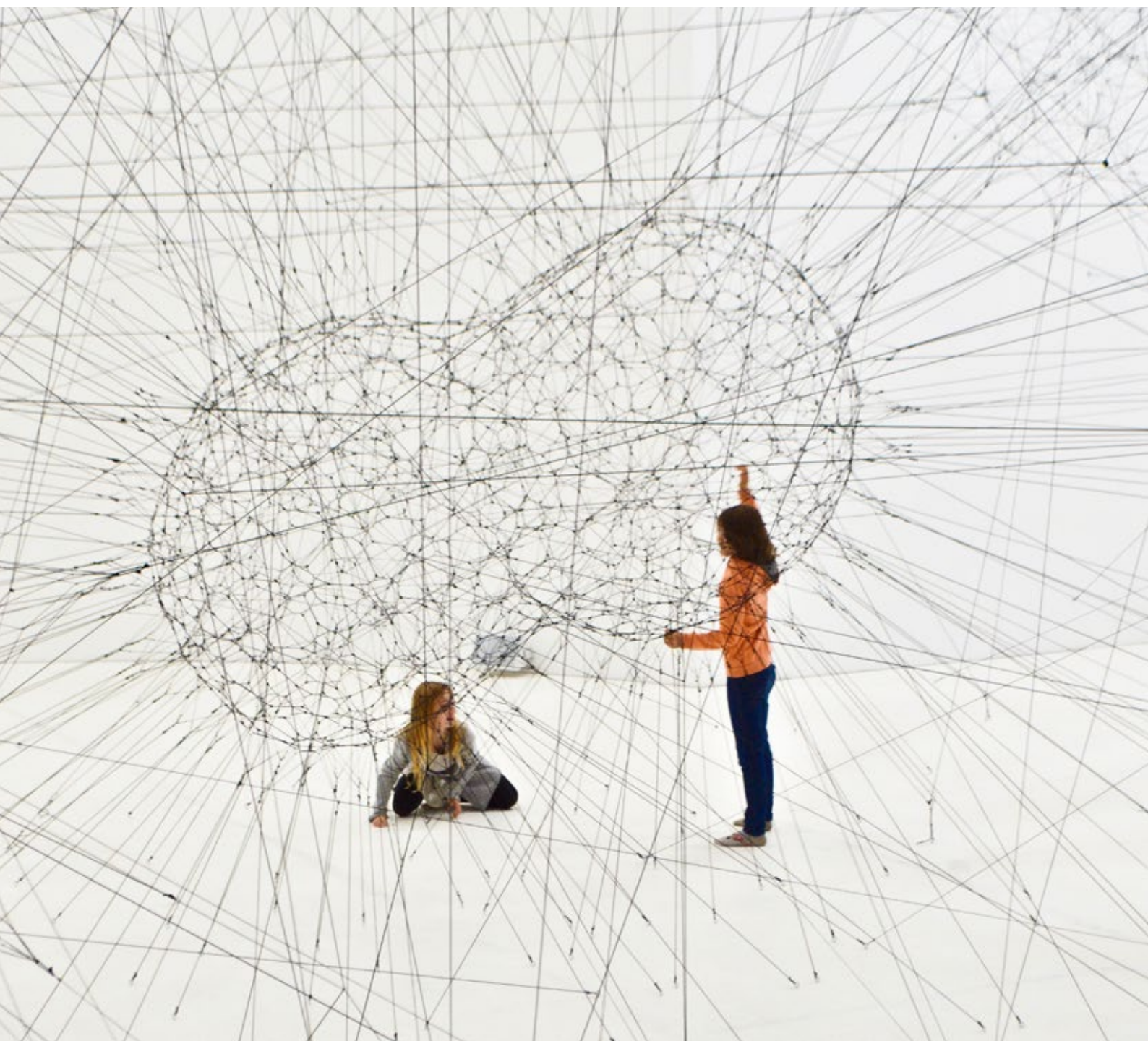


DataEthics.eu

White Paper on Data Ethics in Public Procurement of AI-based Services and Solutions

April 2020



DataEthics.eu 2020

White Paper on Data Ethics in Public Procurement of AI based services and solutions.

Authors:

Gry Hasselbalch
Birgitte Kofod Olsen
Pernille Tranberg

Published by DataEthics.eu

Info@dataethics.eu

CVR 38465724

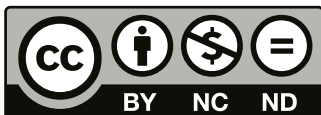
Denmark

Graphic design

PAWs FABRIK

Front page photo

Clarisse Croset, Unsplash



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

ISBN 978-87-972168-0-4

About DataEthics.eu

This white paper has been drafted by DataEthics.eu.

DataEthics is a politically independent non-profit ThinkDoTank founded in Denmark with global reach. We develop independent research, analysis, knowledge exchanges and tools, and raise awareness about data ethics to support a future with human agency at the centre.

We work to ensure the human interest in a fair world of data.

Data systems constantly evolve. From the data of the printing press to the data of the AI agent. Data will either empower or disempower human beings. We trace the data in human society, not the technology.

DATAETHICS

See more: www.dataethics.eu

This white paper was supported by
www.luminategroup.com

Luminate

Building stronger societies

Introduction

This white paper is an independent contribution in support of a general movement among European Union institutions and member states to guide and make strategic use of public procurement of AI, the ultimate goal being to reach shared objectives regarding the adoption and development of trustworthy AI in Europe.

Building on the recommendation of the EU High Level Group on AI to help the public sector foster responsible AI innovation, this white paper supports the development of data ethics and trustworthy AI components as part of the actions on public procurement proposed in the European Commission's 2020 AI white paper and data strategy.

The present white paper covers one pivotal general area of AI adoption in Europe in which a standardized general European approach is of key importance to the adoption of AI that truly respects fundamental rights and values in its governance structures, management systems, and technical, legal and social components.

For a quick read, we recommend the Executive Summary and Recommendations sections.

Contents

1. Background and Purpose	9
1.1. Why We Need a Public Procurement Framework for Trustworthy AI	9
1.2. Procurement of public infrastructure.....	10
1.3. Harnessing the risks	10
1.4. Democratic processes at stake.....	10
1.5. Rights and AI opportunities are complementary.....	11
1.6. Integrating data ethics in AI	11
1.7. Background.....	12
2. Scope and Concepts	13
2.1. Scope.....	13
2.2. Defining Artificial Intelligence (AI).....	13
2.3. A new strategic priority in procurement	14
2.4. Concepts and approaches	15
3. Relevant Legal Frameworks	16
3.1. Fundamental rights and freedoms.....	16
3.2. Equality and societal goals.....	18
3.3. Strategic procurement.....	19
4. Data Ethics in AI Procurement	21
4.1. A common framework.....	21
4.2. Data Ethics Principles.....	21
4.3. Compliance	23
4.4. Accountability	24
4.5. Technical Robustness.....	25
4.6. Sustainability	25
5. Due Diligence in AI Procurement	27
5.1. An adapted due diligence model	27
5.2. The due diligence process in AI procurement	27
6. Preliminary Assessment and Screening	30
6.1. General risk and impact assessments	30
6.2. Stakeholder dialogue	31
7. Preliminary Screening of Potential Suppliers	32
7.1. AI Design	32
7.2. Development and operations (DevOps)	33
7.3. Testing.....	35
8. Contracting	38
8.1. Exclusion Criteria.....	38
8.2. Selection Criteria	38
8.3. Technical specifications	41
8.4. Award criteria	43
9. Contract Performance Conditions	46
10. Contract Implementation	49
11. Recommendations	51
Sources	53

Executive Summary

Technologies are not neutral, neither are choices in the public procurement of AI. The AI systems we deploy today are the systems we will live with tomorrow.

Artificial intelligence is increasingly shaping the opportunities of European citizens and transforming their relationships with governments and public authorities. Today we have the opportunity to responsibly define the way AI will be implemented in the future.

This white paper provides a detailed map for public procurers to choose **AI-based services and solutions that put data ethics, democracy and fundamental rights first**. It is an initial step towards creating a standardized framework for the questions and considerations public procurement processes should include to adopt trustworthy AI and reward the European development of it.

WHY - the quest for a human-centric approach to AI

In late 2019, Ursula Von der Leyen pledged that within her first 100 days in office, she would propose legislation for a coordinated European approach on the human and ethical implications of AI. This white paper suggests integrating data ethics with the strategic approach to public procurement in the EU, making it supplementary to the green and social components that can be implemented as part of the overall legal framework on public procurement.

The social and ethical implications of the absence of a framework based on European values and norms for the adoption of AI in the public sector are already emerging. Automated systems to socially score families, automated systems with poorly-written code erroneously assigning public positions, black box big data analysis systems shared by different state institutions to track citizens etc.

If Europe is to truly innovate in terms of trustworthy AI systems, public institutions must lead the way. If we do not sufficiently address the trustworthy aspects of AI in public procurement, the cost for European societies, individuals and democracy will be high, with changing society forever and citizen's rights potentially being negatively impacted by the systems we choose to implement.

Public procurement of AI technology is not just a matter of choosing between more or less efficient technical tools. It is also a prioritization of interests and values embedded in their design. Trustworthy human-centric AI is an alternative type of innovation that can be developed and thrive in Europe, and the public sector could spur that development with data ethics and principles being hard-wired into public procurement.

WHAT - public procurement as leverage for trustworthy AI

This white paper includes 'data ethics' as a horizontal theme that cuts across the components of trustworthy AI. Data ethics is the responsible and sustainable use of data. It is about doing the right thing for people and society. Data processes should be designed as sustainable solutions benefitting the interests of individual human beings first and foremost. Data ethics is about efforts to create transparency and foreseeability in regard to the social and ethical implications of data processing, and it is about real accountability in governance and management structures. Its goal is actively developing privacy-by-design and privacy-enhancing products and infrastructures and stressing the need always to handle someone

else's personal information in the same way as you wish your own data, or your children's data, were handled.

WHERE - digitized public sector institutions

Strategic, guided public procurement processes could be implemented in public sectors such as:

- Justice and law enforcement
- E-government
- Government to Business (G2B)
- E-democracy
- Education
- Healthcare treatments and services
- Social security services
- Employment.

The process should involve all actors throughout the public procurement process from management, subject-matter experts, designers of systems, data scientists and engineers to civil servants, policy officials and governmental representatives.

HOW - a risk-based and systematic approach to public procurement

Trustworthy AI is possible to achieve by establishing an AI procurement framework that includes data ethics components and applying them within the context of existing legal obligations as well as demands for accountability, technical robustness and sustainability.

This white paper suggests a risk-based approach in public procurement that is aligned with both formalized and applied due diligence processes. It recommends a due diligence process consisting of five phases:

1. **Preliminary risk assessment**

Addressing any adverse impact on human beings or groups of people, their rights and freedoms, on democratic institutions and processes, and on society and the environment.

2. **Preliminary screening of potential suppliers**

This involves screening the market for potential suppliers that possess the necessary skills, competences and organizational structures to fulfil requirements for data ethics components in AI services and solutions.

Data ethical requirements relating to AI should be considered, defined and implemented from the very beginning of the design process. For example:

- When AI systems interact with users directly (e.g. chatbots, virtual assistants) or indirectly (e.g. automated decision-making), they must reveal that they are not human
- AI systems must be traceable, explainable and include stakeholders
- AI systems must avoid bias, be made according to universal design and include procedures for reviews
- Technical robustness should be documented, as should explainability, fair communication and audits.

3. **Contracting**

General **exclusion criteria** should apply when assessing economic operators who have submitted tenders to provide AI-based services and solutions, including past participation in criminal organizations, corruption, fraud, child labour and human trafficking.

Selection criteria should cover relevant specialist technical competences and diverse, multidisciplinary teams that understand the interdependent disciplines that AI covers. In some circumstances, location within EU/EEA should be prioritized. Also, tenderers must guarantee their sub-suppliers comply with the same data ethics standards.

All the **technical specifications** for procurement of AI-based services or solutions should include requirements regarding the methodologies and processes foreseen in the development of the AI-based system or solution.

Award criteria Tenders should be assessed according to a set of economic and quality criteria and a best price-to-quality ratio. The quality criteria should reflect the technical specifications regarding applied standards and management systems for information security, data ethics, environmental aspects, privacy, universal design, etc.

4. **Contract performance conditions** To reach the overall goal of sustainability and respect for fundamental rights, and the specific goal of data ethics in AI-based services and solutions, the contracting authority should include clauses in the contract performance conditions on these issues, along with possible sanctions and documentation requirements.
5. **Contract implementation** A governance structure that includes top management in decision-making processes should support the project structure and identify roles and responsibilities in relation to all phases and levels of the AI project. The supplier should meet the requirements set out in public contracts under five headings: data ethics, legal compliance, accountability, technical robustness and sustainability. To do so, it should set up an organization with insight and overview of the contractual obligations and corresponding work processes.

Further, this white paper recommends (see Recommendations);

- An EU directive on public procurement of AI-based services and solutions for the public sector
- A guiding document on public procurement of AI-based services and solutions for the public sector
- Inclusion of data ethics as a strategic policy priority in the EU Public Procurement Strategy
- A training toolkit on data ethics in procurement of trustworthy AI-based services and solutions
- A handbook on data ethics in procurement of AI-based services and solutions
- An online help desk for public and private sector tenderers.

1. Background and Purpose

1.1. Why We Need a Public Procurement Framework for Trustworthy AI

AI-systems are increasingly becoming part of the socio-technical infrastructures of European societies. Based on complex data collection and analytics, they are shaping the opportunities of European citizens and transforming their relationships with governments and public authorities.

- Yet, we do not have standardized public procurement strategies with ethics and social impact assessments available for their adoption.

We are at an early stage of AI adoption in European societies, and therefore we have the opportunity to shape its direction in a responsible manner. In this context, national governments, public authorities and EU institutions are crucial when paving the way towards trustworthy AI consolidation and making strategic use of public procurement.

Today, European innovation is far from thriving in relation to trustworthy AI. Computer science and engineering students are not trained in the social and ethical implications of AI, most developers are not aware of it, scientists often work in disciplinary silos, businesses struggle to implement and comply with legal requirements without really being creative and innovative with trustworthy AI components. Perhaps most importantly, we do not demand trustworthy AI during public procurement processes.

We are still at an early experimental stage of AI's adoption in society, one in which the technologies and methods required to guarantee trustworthy AI remain scarce and widely debated. This white paper is the first step towards creating a standardized framework for the questions and considerations that public procurement processes should include to adopt trustworthy AI and simultaneously reward European development of it. An EU-wide approach is therefore essential to the creation of a level playing field across the EU for public procurement processes, thereby avoiding market fragmentation.

- Public procurers do not have the awareness, tools and frameworks needed to help them responsibly choose and adopt AI systems.

With digitalization, the public sector is moving towards greater efficiency. Governments and public authorities in Europe are increasingly using AI to streamline and rationalize services and interactions with citizens. AI programs are already in place to assess the risk of violence among adolescents, detect tax evasion, assess student learning patterns, profile unemployed people, find 'irregularities' in citizens' data to detect, for example, social benefit fraud, recognize motion patterns with intelligent video surveillance, and to automate the processing of traffic offences.¹

In ordinary times, ethical choices are difficult. In moments of crisis and emergency, they are even harder. Urgent, rapid decisions are needed to employ the opportunities AI offers to help solve the big problems we are facing. But without proper guidance and methods for assessing social and ethical implications, the cost to European societies and democracy will be high. European societies will transform, and people's rights will be affected by the systems we choose to implement.

¹ Examples from the report "Automating Society Taking Stock of Automated Decision-Making in the EU" (2019), AlgorithmWatch in cooperation with Bertelsmann Stiftung. See more examples here: https://algorithmwatch.org/wp-content/uploads/2019/01/Automating_Society_Report_2019.pdf

- The AI systems we deploy today are the systems we will live with tomorrow.

This white paper taps into a general movement in EU member states to guide and make strategic use of the public procurement of AI to reach shared objectives regarding the general adoption of AI in Europe. It covers one pivotal, overarching area of AI use in Europe, one in which a standardized approach across Europe is key to the implementation of AI that truly respects and safeguards fundamental rights and values in its technical, legal and social components. It will also serve as a guideline for regions around the globe inspired by the European approach to AI.

1.2. Procurement of public infrastructure

Public infrastructure like roads, streets and train tracks points us in specific directions. Public infrastructure enhanced by AI systems also directs our movements, informing or making decisions that shape our opportunities and limits in public and private life. Decisions made by public authorities will increasingly be extended by AI systems with, for example, decisions about individuals' entitlement to welfare services or assessments of educators' and children's performance in the school system. Thus, AI systems will impact people's lives not only in the present, but also in the future.

- Ensuring the stability of roads and train tracks is an infrastructural responsibility, as are the social and ethical implications of AI decision-making systems.

1.3. Harnessing risks

The social and ethical implications of the absence of a guiding framework for the adoption of AI in the public sector based on European values and norms are already emerging in various ways. For example, in Denmark the suggestion to create an automated, data-driven risk assessment system to identify child neglect among vulnerable families, including the assignment of social credit scores to all families, has caused heavy criticism from human rights experts. Nevertheless, this model was later introduced as part of the so-called 'ghetto plan', in which special measures (e.g. harsher punishments for crimes) would be applied in designated 'ghetto' areas in Denmark. In the Netherlands, a group of civil rights initiatives have sparked legal proceedings against the use of a big data analysis system shared by different state institutions to detect citizens who have unlawfully collected public funds. In Spain, an automated system was designed to evaluate how to manage the mobility of teachers as part of educational reform. However, the program's code was so badly written that more and more teachers found themselves assigned to destinations they didn't state in their preferences, which caused general doubt among the population about the reform.²

1.4. Democratic processes at stake

Fundamental relationships between the public institutions, governments and citizens are transforming rapidly. Democracy is built on carefully crafted relationships between citizens and their governments that are embedded in the very fabric of society. As the digital data linking of these relationships, the design and adoption of AI is of crucial importance to the shape of European democracy and society.

² AlgorithmWatch, 2019.

Recent examples of the democratic downsides of digitalization around the globe - from voter manipulation to mass surveillance - compel European societies to revisit fundamental rights and values to ensure that AI is woven into critical infrastructure responsibly and with respect for the continent's ethical frameworks.

- Technologies are not neutral; neither are choices in the public procurement of AI.

The choice to publicly procure one AI program over another is never a mere decision between more or less efficient technical tools. It is also a prioritization of interests and values embedded in their design.

1.5. Rights and AI opportunities are complementary

To innovate is to meet new requirements in new ways, and trustworthy AI is the most innovative investment for Europe.

- Citizens are requesting it. Democracy is requiring it.

Building AI without infringing on fundamental rights and democracy while limiting its risks are not mutually exclusive objectives. Trustworthy, human-centric AI is an alternative type of innovation that can be developed and thrive in Europe. But it does not exist in a bubble. Trustworthy AI needs a technical design and organisational business culture in which people are, for example, in control of and have insight into their data, where processes and systems can be explained and audited, programmers and designers are aware of the social implications of their work, and goals are set for social good.

1.6. Integrating data ethics in AI

AI is data in a specific form. It is developed from data, evolve on the basis of data and act on data in digital environments. Not only is data the building block of AI, it is also a core component of socio-technical infrastructure in contemporary societies.

Data ethics is the responsible and sustainable use of data. It is about doing the right thing for people and society. Data processes should be designed as sustainable solutions benefitting, first and foremost, the interests of individual human beings.

Data ethics is a step beyond mere compliance with personal data protection laws. All data processing therefore respects - as a minimum - the requirements set out in the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights, and those in secondary EU legislation, including the General Data Protection Regulation (GDPR).

Data ethics is also about our understanding and interpretation of legislation. It refers and adheres to the principles and values upon which fundamental rights and personal data protection laws are based. As such, data ethics is about efforts to create transparency and foreseeability with regard to the social and ethical implications of data processing. And it is about real accountability in data management. It pursues a goal of actively developing privacy-by-design and privacy-enhancing products and infrastructures and stresses the constant need to handle someone else's personal information in the same way as you wish your own data, or your children's data, were handled.

1.7. Background

With the EU strategy on AI adopted in 2018, a general governance process was initiated in the EU member states.³ Informed by the High-Level Expert Group on AI's ethics guidelines and policy and investment recommendations for Trustworthy AI, an approach to AI-based on European values and fundamental rights has now been put forward as Europe's answer to AI adoption in the global arena.

In 2019, the new president of the European Commission, Ursula Von der Leyen, pledged that, within her first 100 days in office, she would propose legislation for a coordinated European approach on the human and ethical implications of AI⁴ and accordingly in February 2020, a white paper on AI and a European data strategy was published.⁵

This white paper on public procurement of AI takes its point of departure in the recommendation of the European High-Level Expert Group on AI to support the European public sector in responsible public procurement of AI: 'The public sector can make strategic use of public procurement to foster responsible innovation, as well as steering it towards tackling societal challenges and the development of trustworthy AI solutions.' (Policy & Investment Recommendations, June 2019).

Building on this recommendation, the white paper aims to support the development of the components of **trustworthy AI with an emphasis on data ethics** in Europe to contribute to the following actions proposed in the European Commission's AI white paper and data strategy:

1. To adopt AI programs that will support public procurement of AI systems, and help to transform public procurement processes themselves (white paper)
2. To elaborate a data initiative for public procurement data covering both the EU dimension and the national ones complemented by a procurement data governance framework (data strategy)
3. To facilitate the development of common European standards and requirements for public procurement of data processing services (data strategy)
4. To facilitate the set-up of a cloud services marketplace for EU users from the private and public sector facilitated by the Commission, which will put users (in particular the public sector and SMEs) in the position to select cloud processing, software and platform service offerings that comply with a number of requirements in areas like data protection, security, data portability, energy efficiency and market practice (data strategy).

³ For the EU strategy, see: <https://ec.europa.eu/digital-single-market/en/artificial-intelligence>

⁴ [A Union that strives for more My agenda for Europe, part 3 A Europe Fit for the Digital Age, 2019.](#)

⁵ White Paper on Artificial Intelligence - A European Approach to Excellence and Trust. Brussels, 19.2.2020, COM(2020) 65. A European strategy for data. Brussels, 19.2.2020 COM(2020) 66 final. A European strategy for data. Brussels, 19.2.2020 COM(2020) 66 final.

2. Scope and Concepts

2.1. Scope

The public sector plays an important role in ensuring that AI-based services and solutions are developed, deployed and maintained across the public sector, aligned and compliant with EU standards and regulation, and in pursuit of an ethical application of AI to the benefit of the European citizens.

The aim of this whitepaper is therefore to facilitate a data ethics framework for the procurement of trustworthy AI in the public sector in the EU member states.

By paving the way for AI procurement that is in-line with data ethics, the public sector will have tremendous leverage in relation to innovation and services within AI business and research communities that respect fundamental rights and values and pursue trustworthy solutions.

Thus, the target group of this whitepaper is EU politicians, legislative bodies and institutions, and other relevant stakeholders who are part of the process of developing strategies and regulation within the area of public procurement.

2.2. Defining Artificial Intelligence (AI)

To this day, there has been no commonly accepted, broadly shared definition of the term Artificial Intelligence (AI). Coined in the 1950s, the term has had many meanings and applications, depending on context and application. Recently, AI has re-emerged in public discourse and policymaking as a more generic term to describe a data intensive development of digitalisation.

This white paper builds on the definition of AI presented by the EU High-Level Expert Group on AI:

“ *Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital 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 (...)* High level Expert Group on AI, Ethics Guidelines, p. 36.

In addition, we use the term ‘trustworthy AI’ as defined by the EU High-Level Group on AI to refer to AI solutions that embed legal and ethical principles and requirements in their design and implementation.⁶ There are three layers to this concept, meaning that it should be:

- Lawful, complying with all applicable laws and regulations;
- Ethical, ensuring adherence to ethical principles and values; and
- Robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.⁷

⁶ EU High Level Expert Group on AI, Ethics Guidelines, p. 5

⁷ In this white paper, when using the term robust in specific, we are only referring to technical robustness. We consider social robustness a horizontal component together with data ethics.

Extending these definitions of AI, we would like to emphasize AI as complex socio-technical data systems and processes with different levels of human involvement. We argue that it is this same human involvement in AI data design, organisation, adoption and use that can be influenced and shaped with public procurement. We therefore have included 'data ethics' as a horizontal theme that cuts across the components of trustworthy AI.

DataEthics.eu has developed a set of data ethics principles⁸ that we use as a general framework to assess the data ethics of trustworthy AI:

- The human being at the centre
- Individual data control
- Transparency and explainability
- Accountability
- Equality.

2.3. A new strategic priority in procurement

To facilitate and support this process, this white paper suggests a guiding framework for integrating data ethics considerations and impact assessments into public procurement of AI-based products and services. Its ultimate purpose is to contribute to a wider uptake of *trustworthy AI* by European member states.

The goal is to ensure that future public procurement of AI is aligned with EU Directives on Public Procurement⁹ and the EU Commission's public procurement strategy and supplementary guidelines.¹⁰ Therefore, it addresses two strategic priorities:

- Strategic public procurement: adding 'data ethics' as a new strategic criterion to existing social, environmental and innovative criteria.
- A guiding framework: how to best integrate ethical and legal considerations into public procurement procedures.

We envisage that strategic and guided public procurement of trustworthy AI could be implemented in public administration such as

- Justice and law enforcement
- E-government
- Government to Business (G2B)
- E-democracy
- Education
- Healthcare treatments and services
- Social security services
- Employment.

⁸ <https://dataethics.eu/data-ethics-principles/>

⁹ EU Directive 2014/24/EU on Public Procurement, EU Directive 2014/25/EU on Procurement by entities operating in the water, energy, transport and postal services sectors and EU Directive 2014/23/EU on the award of concession contracts.

¹⁰ EU Public Procurement Strategy (2017), Green Public Procurement initiative (2027) and Innovation Guidelines (2017).

This should involve all actors throughout the public procurement process, from system designers, data scientists and civil servants to policy officials and governmental representatives.

2.4. Concepts and approaches

This whitepaper addresses public procurement of AI-based services and solutions with a fundamental rights approach. Hence, it builds upon the legal framework within EU law on fundamental rights and freedoms as defined in the EU Charter of Fundamental Rights. As it is focused on data ethics, this whitepaper emphasizes the inherent right to dignity, the right to privacy and protection of personal data, non-discrimination and equal opportunities. The use of AI by public administration may impact other fundamental rights such as the freedom of movement, freedom of expression and information, the right to vote, to a fair trial and effective remedies. The whitepaper should thus be seen as only addressing a non-exhaustive list of fundamental rights.

Legal regulation operationalising the fundamental rights to respect for privacy and data protection is found in the EU General Data Protection Regulation (GDPR) and the EU Regulation on the free flow of non-personal data (FDDR). The whitepaper adds a supplementary soft law layer to the legal obligations consisting of the EU High-Level Group on AI's ethics guidelines and policy, and investment recommendations for trustworthy AI¹¹ and the European Commission Whitepaper on Artificial Intelligence.¹² The applied legal framework is described in part 3.1.

Future application of AI-based services and solutions to provide access to public goods and services affects the principles of equal treatment and non-discrimination. As these rights are pivotal for ensuring the overall purpose of the EU, fundamental rights and EU directives pertaining to the area of equality and social goods are included in the white paper. See more in part 3.2.

The suggested framework is modelled upon the existing EU legal framework and strategy for public procurement, cf. part 3.3. to make it feasible to integrate the framework in existing EU structures.

The underlying concept for integrating risk and impact assessments is based on generic models for due diligence as recommended by the EU in relation to non-financial reporting,¹³ in the UN Guiding Principles on Business and Human Rights¹⁴ and the OECD Guidelines for Multinational Enterprises.¹⁵ In setting up the framework for this white paper, applied procurement practices at large corporations have been inspirational. The aim of aligning the suggested framework with existing standards and business practices is to make it recognizable for providers of AI-based services and solutions to the public sector.

¹¹ High Level Expert Group on AI - Ethics Guidelines for Trustworthy AI, (April 2019) and HLEG AI - policy and investment recommendation for trustworthy AI (June 2019).

¹² WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust Brussels, 19.2.2020, COM (2020) 65 final.

¹³ Directive 2014/95/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups of 22 October 2014.

¹⁴ UN Guiding Principles on Business and Human Rights, 2011.

¹⁵ The guidelines were adopted in 2011 and supplemented with OECD Due Diligence Guidance for Responsible Business Conduct in 2018.

3. Relevant Legal Frameworks

For this white paper, we apply EU legal frameworks establishing fundamental principles and values in democratic societies. They encompass fundamental rights and freedoms, especially the right to respect for dignity and privacy, data protection, the protection of vulnerable groups, and social cohesion as they are imperative to thriving and growth in the digital age.

To ensure that the proposed recommendations are feasible, the legal framework and policies for public procurement in the EU have been used as a guide when establishing a proposed structure that is ambitious while also recognizable and realistic.

3.1. Fundamental rights and freedoms

Dignity

The applied concept of giving priority to the interests of human beings in relation to data ethics is embedded in article 1 of the EU Charter of Fundamental Rights and Freedoms, stating that human dignity is inviolable and must be respected and protected.¹⁶

This principle, as it relates to AI, has been applied in accordance with the Council of Europe's Convention on Human Rights and Biomedicine.¹⁷ Article 2 of the Convention stipulates the primacy of the human being and states that 'the interests and welfare of the human being shall prevail over the sole interest of society or science'.

This primacy principle reflects the need to respect human beings both as individuals and as members of the human species, and operationalizes the basic right to dignity of the human being as established in the UN's Universal Declaration of Human Rights, the European Convention on Human Rights and Freedoms¹⁸ and the EU Charter. The preamble of the biomedicine convention thus mentions that the misuse of biology and medicine may lead to acts endangering human dignity. It seems fair to anticipate a similar risk in relation to the misuse of AI-based services used, for example, in automated decision making and individual profiling in the public sector, including individualized risk profiles for tax purposes, job performance in relation professional opportunities, or behaviour or fitness in relation to healthcare treatments.

Privacy and data protection

The EU Charter obliges EU member states to ensure the right to respect for the private life of every individual in article 7 and the right to personal data protection in article 8. The Charter introduced the right to data protection as a separate fundamental right and thereby strengthened the protection of personal data deriving from article 8 of the European Convention on Human Rights and Freedoms.

Secondary legislation in the EU imposes legal obligations on member states to protect

¹⁶ EU Charter for Fundamental Rights, 2000/C 364/01.

¹⁷ Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine, CETS 164, 4 April 1997.

¹⁸ Council of Europe Convention on Human Rights and Fundamental Freedoms, CETS No. 194, 3 September 1953.

personal data. As a result, the GDPR applies to all data processing activities in the public and private sectors, and will be the standard for the development, application and maintenance of services and solutions. With its risk-based approach to data protection and specific principles for processing (article 5), it defines the guarantees for safe and legitimate use of AI, explains the roles and responsibilities for data controllers and data processors (article 24 og 28), including in relation to appropriate security measures reflecting the risk picture (article 32).

An important obligation is put on the data controller in relation to ensuring the implementation of data protection by design and by default, both when determining the means of personal data processing and at the time of the processing itself (article 25).

Moreover, the explicit right of data subjects not to be subject to a decision based solely on automatic processing and profiling if the decision has legal effects concerning the person or similarly significantly affects the person, is of paramount importance in relation to AI. As a result, it introduces a new standard for risk analysis.

The EU directive covering data protection in law enforcement¹⁹ safeguards citizens' fundamental right to data protection whenever personal data is used by criminal law enforcement authorities for law enforcement purposes. It thus plays an important role in the manner in which law enforcement authorities obtain AI-based services and solutions to support decision-making processes as part of the prevention, investigation, detection or prosecution of criminal offences or for the implementation of criminal penalties. In these situations, the directive ensures that protection standards for the personal data of victims, witnesses, and suspects are duly protected and will facilitate cross-border cooperation in the fight against crime and terrorism.

The EU regulation on the free flow of non-personal data (FFDR) is highly relevant in relation to AI as AI-based services and solutions draw on large scale data sets.²⁰ The FFDR regulates mixed data sets of personal and non-personal data. The latter is defined as:

- Data which initially did not relate to an identified or identifiable natural person. E.g. data on maintenance needs for industrial machines or trading data on the financial sector.
- Data which was originally personal data but was later anonymized, e.g. aggregated datasets used for big data analytics or anonymized data used for reporting purposes.

With regards to the second group of non-personal data, the FFDR underscores that anonymization is different than pseudonymization. Pseudonymized data is still considered personal data, as the information can be linked to an individual by combining it with additional data.

The assessment of whether data has been properly anonymized is not a one-size-fits-all operation as it depends on the circumstances of the specific case. As the Commission mentions in the guidelines, if non-personal data can be attributed to a person in any way (and

¹⁹ EU Directive on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, 2016/680 of the European parliament and of the Council, 27 April 2016.

²⁰ EU Regulation on a framework for the free flow of non-personal data in the European Union 2018/1807 of the European parliament and of the Council, 14 November 2018.

hence, make the person identifiable), the data must be considered personal and within the scope of the GDPR.²¹

In a mixed dataset, if the non-personal and personal data are inextricably linked, the GDPR applies to the whole set of mixed data, regardless of the amount of personal data contained in the dataset. In other words, if separating the set of personal data from the non-personal one is impossible, economically inefficient (e.g. if by separating the datasets, its value is compromised), or not technically feasible, the set of data shall be considered inextricably linked and the data protection rights and obligations stemming from the GDPR apply to the whole dataset.

3.2. Equality and societal goals

Equal treatment and non-discrimination

Several examples show that AI can potentially further stereotypes and bias in relation to ethnicity, skin colour, gender and age, highlighting the need for data sources and elements to be thoroughly analysed in AI development processes. Furthermore, it necessitates comprehensive maintenance and auditing of the output and outcome of applied AI-based systems and solutions.

In that light, the foundational principles of equal treatment and non-discrimination stated in the EU Treaty and the EU Charter are of outmost importance in relation to AI procurement.

In the EU Charter, a full section is dedicated to fundamental rights ensuring equality before the law (article 20), non-discrimination (article 21), respect for diversity with regard to culture, religion and linguistics (article 22), equality between men and women (article 23), rights of the child (article 24) and of elderly persons (article 25), and the integration of those with disabilities in order for them to benefit from measures designed to ensure their independence, social and occupational integration, and participation in the life of the community (article 26).

To a large extent, the fundamental rights on equal treatment and non-discrimination have been operationalized through EU directives on discrimination at work on grounds of religion or belief, disability, age or sexual orientation,²² discrimination on grounds of race and ethnic origin,²³ on equal treatment of men and women in employment and occupation,²⁴ and on access to and supply of goods and services.²⁵ The standards and requirements established in the directives are to be interpreted and applied in a digital age and are highly relevant when designing and developing sustainable AI-based systems and solutions.

²¹ Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union, Communication from the commission to the European parliament and the council, COM(2019) 250 final, Brussels, 29.5.2019.

²² Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation

²³ Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin

²⁴ Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation

²⁵ Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services

In the recently-adopted EU Directive on accessibility requirements for products and services, new obligations have been imposed on member states to ensure a more inclusive society and facilitate independence for people with disabilities.²⁶

Also, the concept of universal design that improves the usability of products, environments, programs and services for all people, including those with disabilities, through design was introduced by the UN Convention on the Rights of Persons with Disabilities.²⁷

Common societal goals

EU legislation on public procurement also contains wording relating to the overall goals of society. Thus, the EU directive on public procurement is framed by guiding principles on equal treatment and non-discrimination. Authorities in member states tasked with procuring works, supplies and services are thus expected to comply with these principles, as they derive from the free movement of goods, freedom of establishment and the freedom to provide services.²⁸

The directive furthermore reflects a social purpose in its preamble by listing one of its purposes as helping procurers make better use of public procurement in support of common societal goals.²⁹

3.3. Strategic procurement

High quality and efficiency

The legal framework for public procurement is outlined in the provisions of the EU Treaty on the Functioning of the European Union and in the EU Procurement Directives, adopted in 2014.³⁰ The 2014 Directives were designed to create an open, competitive and well-regulated procurement market. Not only does this guarantee EU citizens high-quality public services, but it also increases efficiency in spending and improves the transparency of procurement processes.

In 2017, as part of the strategy to make the Single Market stronger, the EU Commission published Making Public Procurement Work in and for Europe, a public procurement strategy that defines an overall policy framework and identifies priorities to improve procurement within the EU.³¹

The document establishes six strategic priorities:

- Ensuring wider uptake of strategic public procurement
- Professionalizing public buyers
- Improving access to procurement markets
- Increasing transparency, integrity and better data.

²⁶ EU Directive (EU) 2019/882 on the accessibility requirements for products and services.

²⁷ Signed on 30 March 2007 and in force from 3 May 2008.

²⁸ EU Directive 2014/24/EU of 26 February 2014, preambular provision 1

²⁹ EU Directive 2014/24/EU of 26 February 2014, preambular provision 2.

³⁰ Directive 2014/24/EU on public procurement, Directive 2014/25/EU on procurement by entities operating in the water, energy, transport and postal services sectors, Directive 2014/23/EU on the award of concession contracts

³¹ EU Public Procurement Strategy (2017).

- Boosting the digital transformation of procurement
- Cooperating to procure together.

Innovative, green and social procurement

The priority area aiming at achieving a wider uptake of strategic public procurement is particularly relevant for this white paper. This strategic approach is defined in the document as encompassing innovative, green, and social components.

In this perspective, public procurement should be a tool used to boost innovation in the EU. To move public authorities in this direction, the Commission has published a notice titled ‘Guidance on Innovation Procurement’, establishing the main elements of a relevant policy framework.³²

Besides innovation, strategic procurement should aim to increase the uptake of green and environmentally friendly solutions. Green Public Procurement (GPP) is a voluntary tool that EU member states and public administrations can decide to adopt.³³ GPP is defined as the process whereby public authorities seek to procure goods, services and works with reduced environmental impact throughout their life cycle when compared to standard goods, services and works with the same primary function.

The initiatives implemented by the Commission to facilitate green purchasing encompass:

- Identification of GPP criteria for specific products and services
- A Buying Green Handbook
- A website dedicated to GPP.

With regard to the social aspect of public procurement, European public authorities are urged to buy in a socially responsible manner and purchase ‘fair trade’ products and services that promote sustainable development. To facilitate social procurement purchasing, the Commission has published a specific guide titled ‘Buying social’. Not only does it define a strategy for buying social, but it also sets forth the requirements and specifics for contracting with suppliers and service providers.³⁴

This white paper suggests integrating data ethics into the strategic approach to public procurement in the EU as a supplementary aspect to the communicated innovative, green and social components, implemented as part of the overall legal framework on public procurement.

³² Innovation Guidelines (2017).

³³ Green Public Procurement initiative (2027) and Communication (COM (2008) 400) “Public procurement for a better environment”.

³⁴ Buying Social A Guide to Taking Account of Social Considerations in Public Procurement, European Commission, October 2010.

4. Data Ethics in AI Procurement

4.1. A common framework

To improve the procurement processes of data ethical AI-based services and solutions for the public sector in EU member states, a common framework should be applied, ensuring that an appropriate due diligence process is performed and that relevant requirements are defined and applied.

Due diligence and requirement specifications are instrumental for assessing the quality and security of the solutions offered, and also for revealing the risks inherent to solutions pertaining to fundamental rights, especially privacy, data protection, and non-discrimination rights.

Moreover, a generic set of requirements and specifications may be used when designing and qualifying development processes, i.e. where public procurement is aimed at purchasing new AI-based solutions and systems that are not yet available on the market.

This white paper suggests an AI procurement framework that is based on a set of data ethics principles that are to be applied within the context of legal obligations and demands for accountability, technical robustness and sustainability.



Figure 1 An AI procurement framework

4.2. Data Ethics Principles

This white paper recommends a trustworthy AI framework for public procurement that horizontally incorporates a set of data ethics principles reflecting individual agency and individual data control as core elements for maintaining human dignity in a digitized age.

The rationale is that every individual should be in control of his or her personal data and be empowered by access to data. A person's self-determination should be prioritized in all data processes and each person should be actively involved in the recording of his/her data. Individuals should have primary control over the usage of their data, the context in which said data is processed and how it is activated.³⁵

³⁵ See DataEthics.eu data ethics principles.

These data ethics principles also aim to ensure an effective European data sharing infrastructure that makes it possible to explore and use data for common societal goals and to boost sustainable growth. However, this objective cannot be achieved without ethical and social risks if the technical, organizational and market conditions are not developed in Europe and supported within a European institutional framework.

The applied principles are aligned with the requirements for trustworthy AI suggested in the High-Level Expert Group on AI's ethics guidelines, the EU White Paper on AI, and the considerations laid out in section 3 above. Hence, they cover:

Human primacy

Human interests always prevail over institutional and commercial interests. People are not computer processes or pieces of software, but unique, with empathy, self-determination, unpredictability, intuition and creativity. AI-based services and solutions should support human autonomy and enable them to prosper in democratic, sustainable and robust societies.

Universal design

Everyone should benefit from products, environments, programs and services based on AI, regardless of their abilities. AI design should not exclude or hinder people from thriving in digitized societies on equal terms but allow for all people to use AI-based products and services, irrespective of their age, gender, abilities or characteristics. Universal design should be used to empower people with disabilities in particular and enable equitable access and active participation.

Transparency

AI-based data processing activities and automated decisions, including profiling, must make sense for individuals. They should not undermine human autonomy but provide individuals with the knowledge and tools required to comprehend and interact with AI-based services and solutions, and to support them in making informed choices and decisions. The purpose and outcomes of data processing must therefore be clearly understood by each individual in terms of comprehending risks and any social, ethical and societal consequences.

Traceability

AI-based services and solutions should be developed and applied in a way that makes it possible to verify and document data sources, data sets and categories, the algorithm used and its variables, and the processes that lead to automated decision making. Therefore, errors and adverse impact on individuals and society can be identified and prevented.

Explainability

The technical processes and related human decisions in AI-based services and solutions should always be explainable. If the AI system or solution produces results or decisions based on correlations, a requirement for causality should be provided for by ensuring human intervention in the decision-making process. The demand for explainability is of vital importance both when the AI-based system or solution has significant impact on people's lives and when it influences and shapes organizational decision-making processes, e.g. in relation

to procurement, design choices and when defining the purpose of deploying the system/solution.

Fair communication

AI-based services and solutions should identify themselves as non-human by informing human users that they are interacting with an AI system/solution. Such information should include notice of the level of accuracy, the risk of adverse impact, and limitations to AI practitioners and end-users. Individuals should also be informed of how they can opt out of AI interactions and demand that human interaction is provided.

Data sharing spaces based on personal data control

AI is data-hungry. It evolves, learns, predicts and decides on data. As such, many ethical implications of AI regard data ethics by their very nature. One overarching ethical challenge is the use of people's data for AI innovation. To ensure a new paradigm in which data asymmetries between institutions, businesses and individuals are limited, a European data infrastructure must be based on individual empowerment and personal data control not only with imposed requirements, but also naturally evolving with the help of motivation and reward systems.

A current movement in technology and business development is addressing this type of individually-controlled data sharing infrastructure with the creation of personal data management systems and services. These are interoperable services that enable individuals to share data and either donate or activate their data for personal benefits (such as personalized finance or medicine) while being in control of the use of their data. Examples can be found in the **Mydata.org** network of entrepreneurs, activists, academics, corporations, public agencies and developers, and in technological initiatives such as **Solid**.

4.3. Compliance

AI-based services and solutions of relevance from a public procurement perspective must comply with the legal obligations of all applicable laws and regulations. This is the component of trustworthy AI that the AI HLEG has referred to as 'lawful AI'. This includes, but is not limited to, legislation on privacy and data protection, especially the GDPR, and supplementary national legislation as well as sector-specific regulations.

Legal regulation on state surveillance and profiling as part of crime investigation and prevention, such as control and monitoring schemes in relation to tax and social benefit fraud, is also highly relevant to AI procurement.

Access to the labour market or to health and social services illustrates areas where AI-based services and solutions are introduced to facilitate job matching, healthcare visits or education by screening and profiling users or applicants. The procurement and application of such AI-based services and solutions calls for thorough compliance-oriented analyses of anti-discrimination laws and public administration law.

Similarly, legislation that furthers freedom of movement, freedom of association and assembly, and participation in politics and cultural life must be complied with when procuring AI-based systems and solutions for smart city programmes, among other things.

The application of conformity assessments for high-risk AI uses as stated in the EU's AI white paper (February 2020) are particularly important when procuring AI services and products from outside the EU, in which case it is pivotal to ensure that European legal and ethical standards are still met.

Other areas of legal concern in relation to AI cover the right to vote and to run for election, the right to good administration and access to public documents. The right to file complaints on decisions made by public authorities or processes failing to comply with public administration law is another example that must be taken into consideration when procuring AI-based services and solutions.

Linked to this area is the need for compliance analyses that unveil potential negative impact on the fundamental principles of democracy, justice and the rule of law, e.g. through the application of AI-based services and solutions undermining democratic processes, the plurality of values regarding individual life choices and the basic principle of equality.

A risk-based approach to the procurement of AI is crucial. Areas in which the application of AI carries particular high risks and potential for adverse consequences requires particular attention, such as the tracking and profiling of citizens, citizen scoring, biometric recognition (e.g. facial recognition), predictive policing, and the monitoring and assessment of children.

Compliance with ethical standards

An additional 'ethical standards compliance' layer may be considered, taking internationally established ethical standards in the research and scientific fields into account. These are already formally applied in the form of ethics audits and assessments within the European funding programme Horizon2020. A few relevant examples are:

[The Helsinki declaration](#)

[The Oviedo Convention \(the 'bioethics convention'\)](#)

See also: [Horizon2020 ethics guidance framework](#)

4.4. Accountability

Closely linked to the demand for legal compliance is the fulfilment of accountability principles. Along with the GDPR and the introduction of a risk-based approach to data protection and the obligation to set up organisational and technical measures, accountability is an important means to achieve and document compliance.

Data protection management systems and defined governance structures, including roles and responsibilities, internal policies, procedures, guidelines, training, communication, monitoring and control mechanisms, have - in practice - become a prerequisite for ensuring the maintenance and review of compliant data-processing activities.

In relation to AI, an accountability approach ensures a reflective, reasonable and systematic use and protection of personal data by forming an integral part of all aspects of data processing, and efforts can be made to reduce the risks for the individual and to mitigate social and ethical implications. It also contributes to the procurement process by highlighting the need to

evaluate whether subcontractors have incorporated accountability measures throughout their organization.

4.5. Technical Robustness

Technical robustness in AI-based services and solutions is imperative for ensuring an effective level of information security and cyber security.

Technical robustness encompasses adequate protection of the integrity and confidentiality of personal data and provides a protective layer against unauthorized access to data; alteration, loss or destruction of data; and unauthorized disclosure of data.

Resilience to hostile attacks such as hacking, malware and spyware, espionage and destruction of infrastructure should form part of effective security systems and measures to protect AI-based services and solutions.

Other technical measures should include backup plans, e.g. for switching from an algorithm-based procedure to human interaction, and mechanisms that ensure that unintended consequences and errors are minimized. Technical measures ensuring correct, accurate decisions, predictions, and recommendations are other safeguards that will help make AI-based services and solutions robust as they help prevent harm to human beings and society.

Similarly, measures that ensure reliable and reproducible results contribute to properly functioning systems and enhance transparency and explainability.

4.6. Sustainability

Preventing harm and ensuring fairness are core aspects of sustainable personal data processing. They address both environmental and social concerns with the goal of benefitting all human beings, including future generations. The two concerns may be intertwined, as is the case in the UN Sustainable Development Goals, by tackling a pressing social need in an environmentally friendly way, for example.

Environmental well-being could also be pursued by integrating the examination of resource and energy consumption in the development and deployment phase of the AI-based system or solution.

Environmental Impact of Data Processing

Algorithmic decision-making has immense impact on the environment due to the extensive data processing and analysis it requires. As an example, the consumption of energy to run a Bitcoin network is equivalent to the energy needs of Cambridge University for 360 years. <https://www.cbeci.org/comparisons>

A goal of social well-being would encourage greater focus on the risks generated by simulating sociality, relationships and emotional attachment when interacting with the chatbots, robots or virtual reality avatars of an AI-based system or solution.³⁶ Thereby, it potentially impacts

³⁶ Examples are listed in the EU High Level Group on AI's, ethics guidelines p. 19.

the physical and mental well-being of humans and may alter socio-cultural practices. Other societal risks are related to its impact on democratic institutions and AI-based decision-making processes.

Assessments of the sustainability of data processing in AI-based services and solutions should thus encompass, as a minimum, the identification of social and environmental risks and adequate measures to minimize intentional or unintentional harm in the short, medium and long term. As society changes, the AI model must be seen as a dynamic tool that requires continuous control with applied variables. Variables that appeared relevant and essential at the time of development may lead to error synchronization, because of changes in behavioural patterns, perceptions and other contextual factors, and produce harmful results. The assessment of sustainability should thus be integrated in regular and systematic application and monitoring processes.

Data ethical trade

During public procurement processes, public institutions should be able to purchase goods which make a special contribution to sustainable development that is ethical in terms of its data. The AI white paper issued by the EU suggests the establishment of a 'voluntary labelling scheme' for low-risk AI applications that would allow users to easily recognize AI-enabled products and services that are in compliance with certain objective, standardized EU-wide benchmarks, going beyond the normally applicable legal obligations. The criteria of such labelling schemes could be incorporated as considerations in, for example tender, specifications (similar to how 'ethical trade' is handled in the ['Buying Social' guide](#)).

5. Due Diligence in AI Procurement

5.1. An adapted due diligence model

As part of a strengthened focus on good governance and sustainability across sectors, concepts and models for due diligence processes have emerged that effectively help organizations to gain insight into their potential negative impact on the environment, human beings, society and the economy. They can be grouped into three main strands:

One recommendation for due diligence processes is embedded in the EU directive regarding the disclosure of non-financial and diversity information and its supplementary guidelines.³⁷ It reflects an acknowledged and widely-used method developed by the UN in relation to the responsibility of corporations to respect human rights,³⁸ and a method presented by the OECD with regard to overall sustainability in the business sector, including human rights, labour rights, the environment, consumer rights, bribery and taxation. The OECD guidelines have led to standards being set for due diligence processes in general and in specific sectors, including institutional investors.³⁹ Most recently, a tool was launched to guide the planning of due diligence processes in connection with the UN Sustainable Development Goals.⁴⁰

In parallel, large corporations have established formalized procurement processes that include third-party risk management programmes. Such programmes are designed to diligently identify and handle risks as part of the procuring, contracting and implementation phases.

As a third strand, public sector procurement has been aligned to the principles and requirements established in EU directives on procurement.

This whitepaper is inspired by all three strands and uses an adapted model as a basis for the following sections, sharing its logic and concepts with methodologies developed in international fora and in the EU. This model may easily be aligned with the formalized procurement structures of contracting public sector entities and integrated with existing practical structures pertaining to tenderers in the business sector.

5.2. The due diligence process in AI procurement

The due diligence process in relation to the public procurement of AI-based systems and solutions improves transparency in the planned AI model and its impact. It helps ensure that all relevant risks and impacts in relation to data ethics are identified and handled, and that the development, deployment and maintenance of the AI model accommodates the needs and requirements of a trustworthy AI-based system or solution. As such, it must span the whole lifecycle of the AI-based system or solution.

³⁷ Directive 2014/95/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups of 22 October 2014 and Communication from the Commission - Guidelines on non-financial reporting (methodology for reporting non-financial information) 2017/C 215/01.

³⁸ UN Guiding Principles on Business and Human Rights, 2011.

³⁹ OECD Due Diligence Guidance for Responsible Business Conduct, 2018.

⁴⁰ See the SDG Compass - The guide for business action on the SDGs.

The due diligence process consists of six phases:

1. Preliminary risk assessment
2. Preliminary screening of potential suppliers
3. Contracting
4. Contract performance conditions
5. Contract implementation
6. Transition

The process thus encompasses an initial phase in which a preliminary risk assessment is completed with regard to the scope/matter the AI-based system or solution is designed to handle (e.g. automated or assisted decision-making for triage in healthcare services).

These assessments should be performed prior to screening potential suppliers or tenderers and before defining requirements and specifications. In this way, the due diligence process informs and facilitates a sharper focus on the demands and specific requirements of a planned AI-based system or solution. The results from this phase may also be used in the process of drafting the tender.

In phase 2, potential suppliers are evaluated, and their performance capabilities are assessed against the expected level of skills, competences, methodologies, governance and procedures for development, training and validating data, testing environments, maintenance and audit functions. This may include information from relevant stakeholders, e.g. academia, groups of impacted citizens, or local committees.

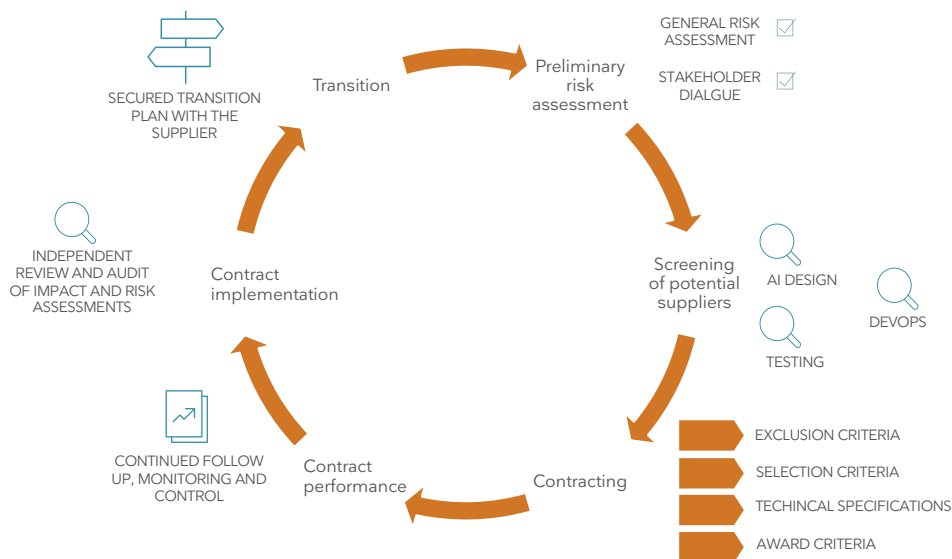


Figure 2 Due diligence process in AI procurement

In the contracting phase, a set of criteria should be adopted that guides the selection of relevant tenderers, allowing for the exclusion of others. Technical specifications help assess, in depth, whether the tenderers can accommodate the integration of data ethics in their development, deployment and maintenance operations. Well-defined award criteria will assist in the final selection of a supplier.

The fourth phase of the due diligence process focuses on the conditions for contract performance. Contract clauses on data ethics should include relevant management and control systems or guidelines to be followed by the contracting party to ensure - as a minimum - that requirements in relation to data ethics, legal compliance, accountability, technical robustness and sustainability are met.

During the fifth phase, the contract should be implemented with a focus on processes and tools that ensure that the contracting party can identify, evaluate and report on progress, including on any risks or incidents, to the public contractor.

In the final phase, the contract is concluded, and the AI-based system or solution securely transmitted to the public contractor. This phase is not covered by the white paper.

The content and requirements set up within the first five phases are described in the following sections.

6. Preliminary Assessment and Screening

6.1. General risk and impact assessments

This white paper suggests that risk assessments be carried out as part of the first phase of the AI procurement due diligence process. The aim is to provide an initial assessment of the risk and potential impact of a planned AI project.

Impact and risk assessments should uncover potential negative outcomes, risks and threats linked to the design, development, deployment and maintenance of a given AI-based system or solution. It should thus cover the expected results from an AI-based decision-making process or interaction.

Preliminary assessments could encompass five sub-categories:

- **Data ethics impact assessment**

A data ethics assessment should address the potential adverse impact on individuals or groups of individuals and democratic societies, and the technical and organizational measures put in place to reduce and monitor such impact. Part 5 and 6 of this white paper contain components that could be transformed into a data ethics impact assessment tool.

- **Legal compliance assessment**

Data protection impact assessments (DPIA) should be performed in accordance with the GDPR to unveil risks to fundamental rights and freedoms. If the planned AI-based service and solution is to be deployed in areas governed by national legislation, e.g. education, the workplace, employment, healthcare or social services, patient rights etc., a compliance assessment of relevant legislation should be performed.

- **Accountability impact assessment**

An accountability assessment is necessary if the expected risks and impacts of the AI-based service or solution are linked to governance structures and managements systems. The assessment should focus on organizational measures and cover adequacy and effectiveness in relation to designing, developing, deploying and maintaining the planned AI project.

- **Security risk assessment**

An assessment of security measures and the security level should focus on risks and negative impact stemming from a lack of technical robustness, including the integrity and confidentiality of the data to be processed by the AI-based service or solution, accessibility to data, cyber security measures, business continuity plans and incident response plans.

- **Social and environmental impact assessment**

If a planned AI-based service or solution is expected to impact democratic institutions and processes, or have excessive impact on the environment, an assessment should be performed to identify the extent, severity and likelihood of such impact.

Whether the assessments should be performed independently or in combination depends on the reach of the AI-based service or solution itself.⁴¹ If it will impact multiple arenas of

⁴¹ See an example of extending the GDPR DPIA to cover good governance and good process in Swee Leng Harris,

human activity or behaviour, and entails extensive impact on society and the environment, the assessment should include all of the categories above. The same is true if sensitive information about citizens is collected and processed in IT systems and networks in which security measures do not provide appropriate and effective protection. If the effect on citizens is insignificant or minor, a preliminary assessment could concentrate on governance and accountability measures in combination with a data ethics impact assessment.

Assessment Tools and Methodologies

There is a need to develop standardized risk assessment methodologies and tools that cover the areas mentioned in this section. Several assessment tools for AI and algorithmic impacts have already been developed by various organizations. For example:

‘Algorithmic Impact Assessments’ by AINOW

<https://ainowinstitute.org/aiareport2018.pdf>

‘Artificial Intelligence Impact Assessment’ by ECP

<https://ecp.nl/wp-content/uploads/2019/01/Artificial-Intelligence-Impact-Assessment-English.pdf>

‘Algorithmic Impact Assessment’ by the Canadian government

<https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai/algorithmic-impact-assessment.html>

‘The Assessment List’ by the EU HLEG on AI

<https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/2>

In addition, guidelines and methodologies exist for assessing the impact of research and science which may be used as a basis for developing such tools. For example, the European Commission has published a comprehensive guide for the assessment of research projects in relation to ethics and data protection:

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf (2018)

6.2. Stakeholder dialogue

As part of the due diligence process, consultations with stakeholders and rights holders should be performed in order to expose potential negative impact on or harm to people or groups of people due to their age, gender, ethnicity, sexual orientation, disability or other characteristics. Stakeholder dialogue is especially relevant where gaps in information exist.

The inclusion of other stakeholders such as suppliers, industrial federations, authorities, academia and non-governmental organizations (e.g. those for consumers, citizens or patients) in active dialogue prior to the procurement process would provide useful insight into risk prediction and risk handling, supply chain mechanisms, and marked maturity in relation to planned AI-based services and solutions.

Data Protection Impact Assessments as rule of law governance mechanisms, Data and Policy, vol. 2, 2020, e2, Cambridge University Press.

7. Preliminary Screening of Potential Suppliers

The following should be considered as part of the second phase of the due diligence process. The purpose is to screen the market for potential suppliers that possess the necessary skills and competences and will implement the proper procedures relating to data ethics in AI products and services.

7.1. AI Design

The development of an AI system should focus on delivering a product that is both technically robust and ethical. To ensure this, those involved in designing the AI-based system or solution should follow methodology composed of specific key activities.

Data ethics requirements for AI should be considered, defined and implemented from the very start of its development, meaning in the early phases of designing and establishing its architecture. Implementing ethical considerations from day one will make it possible to develop an efficient yet cost-effective solution.

The preliminary screening of suppliers of AI-based services or solutions should focus on the implementation of organizational and technical measures that provide insight into their general methodologies, governance and structures.

The fundamental concept in this phase is transparency, which should include the following requirements and criteria:

Self-identification of an AI system

When AI systems interact with users directly (e.g. chatbots, virtual assistants) or indirectly (e.g. automated decision-making), they must reveal that they are not human. This ensures that individuals can make informed decisions as to whether they wish to use the provided service or product. Moreover, users must always be given the possibility to request the intervention of a human being.

Traceability

While designing AI solutions, developers must present a clear record of the systems, algorithm training methods, models, etc. used, as well as the decisions made. It should therefore be possible to keep track of the decisions made by the AI algorithm itself. Traceability facilitates the identification of mistakes and makes their correction easier.

Explainability

Explainability refers to the decisions made by the AI system. The choices made should be comprehensible to individuals interacting with the AI. However, it should also be possible for programmers and coders to explain the combination and weighing of data during the training of the algorithm and in the final application of the data. Hence, an explainable AI can describe the choices made and the outcomes of those decisions. Explainability - as with traceability - should be kept in mind from the beginning of the AI system design process. This is possible, for instance, by selecting the simplest and most interpretable model for the AI system, transparency of the business model selected, etc.

Inclusion of stakeholders

Formalized processes for dialogue with individuals, groups of individuals or their representatives whose life and fundamental rights are affected by the AI can provide insight about adverse impact and risks. Furthermore, inclusion of other relevant stakeholders can inform procurers of the market's readiness to engage in the sustainable development of AI-

based services and solutions.

RECOMMENDATION: Preliminary screening - AI design

Requirement	Criteria
Self-identification	<ul style="list-style-type: none"> ▪ Notification of non-human interaction ▪ Documentation for easy-to-understand and accessible notifications
Traceability	<ul style="list-style-type: none"> ▪ Adoption of procedures for documenting the development and maintenance of AI-based services and solutions ▪ Demonstration of records
Explainability	<ul style="list-style-type: none"> ▪ Explanation of logic and mathematics behind algorithms, choices and deselections, outputs and outcomes
Inclusion of stakeholders	<ul style="list-style-type: none"> ▪ Rights holders impacted by planned AI ▪ Researchers and market representatives

7.2. Development and operations (DevOps)

The procurement of AI-based solutions and systems for the public sector should ensure that data ethics have been integrated in both development and planned operations with potential suppliers. Hence, a preliminary screening should focus on business decision makers, and the involvement of relevant experts and data scientists.

On the technical side, the tasks of programmers and coders should be transparent, and they should account for their analyses and be able to explain choices and decisions relating to data sources, variables in the algorithm, or operation and maintenance flows in the programming of the system. This should also include a reflection on and explanation of the consequences of leaving out certain sources, variables or data flows. Ethical considerations should be integrated in each phase of development, including when writing, analysing and reviewing the code that forms the basis of that AI solution.

This presupposes that the DevOps team is trained in data protection legislation and information security requirements. Moreover, team members should possess knowledge about data ethics as it relates to their jobs, including that they learn how to identify a code's vulnerabilities and risks as they relate to the fundamental rights and freedoms of individuals, and have access to tools or methods to eliminate or minimize them.

Data ethics within coding should be ensured by requesting that suppliers fulfil the following criteria:

Avoid bias

Trustworthy AI systems should not have a discriminatory effect, directly or indirectly. Identified bias should be considered as soon as it is discovered and removed if deemed discriminatory. Potential suppliers should demonstrate how procedures and checklists support this process.

To expedite the detection of bias and the risk of discrimination, it can be beneficial to ensure diversity in DevOps teams. This may also be achieved by including potentially-impacted stakeholders in relevant phases. Furthermore, it is important to appoint a person who is responsible for overseeing the process and who will examine the purpose of each development phase, its limits, and check that the set requirements are met by the AI.

Universal design

In designing AI-based services and solutions, the concept of universal design should be considered and applied to the widest extent possible. Enabling accessibility and participation through suitable user interfaces that allows minority groups such as people with disabilities, cultural or linguistic minorities, children or elderly people to benefit from AI-based solutions would thus further inclusion and diversity in society.

To educate developers about specific needs and preferences among minority groups, dialogue with stakeholders, representatives from relevant minority groups and NGOs, academia, civil servants and other practitioners, should form part of the development process. Similarly, stakeholders could be included in the testing of developed systems and solutions.

Programming and code review

To achieve high ethical standards in an AI, a code review phase is essential. This consists of a manual revision of the code performed by fellow programmers with the aim of detecting and correcting mistakes. Reviewing code can ensure that there is no leakage or excessive processing of personal data. The reviewer should have the goal of identifying where data is stored and ensure that an adequate level of protection is in place.

RECOMMENDATION: Preliminary screening - development and operations

Requirement	Criteria
Avoid bias	<ul style="list-style-type: none">▪ Training programme for DevOps teams▪ Methodologies for identifying bias with a discriminatory effect▪ Procedures for detecting bias and corrective measures
Universal design	<ul style="list-style-type: none">▪ Inclusion of stakeholders▪ Testing accessibility
Programming and code review	<ul style="list-style-type: none">▪ Procedures for internal and external reviews

7.3. Testing

In this phase, the AI system or solution needs to be assessed to ensure that it meets the stated requirements and all checks and balances are in place before it's launched. Moreover, programmers need to test the AI to detect any technical vulnerabilities.

Technical robustness

In the testing phase, security measures need to be assessed to evaluate if the correct level of protection to individuals has been achieved. Therefore, the resilience of the AI system to potential attacks and data leaks need to be tested.. This should be done through the several types of scans, including but not limited to security, penetration, fuzz and dynamic testing.

It should be documented that the applied security measures are adequate and appropriate, and that the resilience of the AI-based service or solution has been tested and meets the necessary level of security.

Traceability

All relevant tests of the AI-based system or solution should be documented. This should include descriptions and a track record of methods used to test and validate the program and the applied algorithm.

Testing and validation should include training data and the scenario tested on such data, and the data used as part of its real-world application.

Explainability

The output of an AI-based system or solution should be understood by human experts. This may include the application of explainable artificial intelligence (xAI) methodology. Explanations should cover the correlation of data, including misfit risk. If possible, the explanation should also encompass the underlying rationale for causality between the input

data processed by machine learning and the expected decision or proposal put forward by the AI-based system or solution.

In this light, systems based on 'black box' concepts should not be seen as fulfilling data ethics requirements.

Fair communication

The supplier should make use of open source methodologies, programming, coding and testing tools. Testing procedures and results should be made available and communicated in a transparent manner.

Maintenance

Monitoring mechanisms should form an integral part of maintenance procedures for AI-based services and solutions regarding security, accuracy, transparency, reliability, correlation and causality. A monitoring routine could consist of three steps: 1) testing, assessing and evaluating, 2) updating and 3) auditing to be performed annually or biannually depending on the risk picture related to the processing activities and purpose of the AI-based system or solution.

Audit

Apart from an audit of the quality and technical robustness of the applied AI system, the potential supplier should demonstrate that output and impact assessments are audited in relation to sustainability, covering the relative environmental footprint, social inequality, exclusion or marginalization of certain groups, and other ethical implications.

The audit should be external and could be based on:

- The auditors' observations and investigations
- Self-assessment by the tenderers
- Documentation, provided by the tenderer, of established measures, the inclusion of stakeholders and impact assessments performed.

RECOMMENDATION: Preliminary screening – testing

Requirement	Criteria
Technical Robustness	<ul style="list-style-type: none"> ▪ Documentation of appropriate security measures ▪ Resilience tests
Traceability	<ul style="list-style-type: none"> ▪ Documentation of testing and validating the algorithm ▪ Validation of training data ▪ Procedure for testing and validation
Explainability	<ul style="list-style-type: none"> ▪ Explanation of correlation ▪ Explanation of the underlying rationale for causality ▪ Applied standards for xAI
Fair communication	<ul style="list-style-type: none"> ▪ Open source ▪ Transparent testing procedures
Maintenance	<p>Testing and evaluation of:</p> <ul style="list-style-type: none"> ▪ Security ▪ Transparency ▪ Quality (accuracy, reliability) ▪ Causality
Audit	<p>Checking for:</p> <ul style="list-style-type: none"> ▪ Discriminatory bias and appropriate mitigating measures ▪ Stakeholder inclusion ▪ Programming and code review ▪ Fair communication ▪ Maintenance management ▪ Environmental impact ▪ Social impact ▪ Ethical impact

8. Contracting

The existing public procurement framework contains concepts, procedures and requirements that may be used as leverage to ensure the integration of data ethics throughout the procurement process.

The following sections illustrate how data ethics principles may be integrated in all phases of the procurement process, adding supplementary elements to already-existing procedures.

8.1. Exclusion Criteria

General exclusion criteria should be applied when assessing economic operators submitting tenders for AI-based services and solutions, including participation in criminal organizations, corruption, fraud, and child labour and human trafficking violations.

Supplementary data ethics exclusion criteria should encompass violations established by national, regional or international courts, and EU or Council of Europe institutions, such as:

- Profiling or assigning a score to citizens or consumers in violation of fundamental rights
- Large-scale identification and tracking of individuals without a specific purpose
- Development and deployment of lethal autonomous weapon systems.

RECOMMENDATION: Procurement process - exclusion criteria

Exclusion criteria	Content
Unlawful exploitation of personal data	<ul style="list-style-type: none">▪ Profiling or social scoring▪ Identification and tracking▪ Lethal autonomous weapon systems
Hostile use of personal data	<ul style="list-style-type: none">▪ Abusing the fundamental right to dignity

8.2. Selection Criteria

To be selected, the tenderer should meet a set of requirements linked to their technical and professional abilities, including human resources and technical capacities and experience. The selection criteria may also include requirements regarding technical facilities, the use of sub-suppliers and verification mechanisms applied to ensure the quality, compliance and security of the sub-suppliers, and references to previously-fulfilled contracts.

Diverse, multidisciplinary teams

A vital component for ensuring that data ethics considerations become an integral part of the procurement process is to assess whether the tenderer has or will have a diverse, multidisciplinary team that understands the interdependent disciplines that AI covers.

A variety of skills and experience is needed to cover all aspects of the AI life cycle. As part of preliminary research and in the design phase, the tenderer should show that it has access to specialized knowledge on how to identify relevant data sources and data elements, while also ensuring compliance with data protection regulation.

When initiating the design of the AI system, the tenderer should make use of team members with expertise in AI systems, data analytics and engineering, model development (e.g. deep learning), model/information visualisation and experience with relevant and reliable AI methods such as text analysis, sentiment analysis, content categorization, process analysis, or augmented government. Also, an agile process to ensure compliance and risk testing during the development, coding and testing phases is imperative. This will typically require specific knowledge and skills relating to fundamental rights, data protection legislation, information security and cyber security.

Deploying and implementing AI-systems in practice requires technical skills (e.g. user interfaces), as well as organizational skills within governance, change management, monitoring and audit programmes, training and awareness-raising, just to name a few.

Technical facilities

Companies who submit a tender should meet specific requirements regarding access to their relevant specialist/technical facilities. The development environment should include state-of-the-art facilities, with methodologies, software, programs, devices and processes for the development team, and relevant tools for integrating data ethics throughout the development process. Similar facilities and tools should be available during testing processes.

The tenderer should also document the installation and maintenance of adequate and robust physical and IT security measures within their facilities and on applied technical equipment, including hardware, software, servers, cloud solutions and networks, and during data training (ML), and the transmission and storage of data. An effective access management system should be in place along with mechanisms set up to track unauthorized access to data or the disclosure, alteration or disappearance of data.

In certain circumstances, tenderers who will locate their technical facilities within the EU/EEA should be given priority. This could be relevant in relation to the processing of sensitive data in combination with critical infrastructure, e.g. structures serving a vital societal interest in relation to energy, health, security, food, transportation and the economy.⁴²

Sub-suppliers

If the tenderer plans to use sub-suppliers, the tenderer should specify which part of the contract will be performed by each sub-supplier, and specify their geographical location within or outside the EU/EEA, technical facilities, diversity and skills, efficiency, experience and reliability in relation to developing and/or applying AI/ML methodologies and tools.

The tenderer should guarantee that sub-suppliers have not been involved in any abuses of fundamental rights and freedoms or complicit in violations covered by the exclusion criteria.

⁴² EU Directive COUNCIL DIRECTIVE 2008/114/EC of 8 December 2008 on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection.

This presupposes that risk assessments have been carried out by the tenderer in relation to data ethics, fundamental rights, accountability, security and sustainability (social and environment) and can be documented.

Tenderers should specify how sub-suppliers will be administered during the processes covered by the procurement contract and what monitoring/auditing mechanisms will be set up in order to monitor and review the quality of the work performed by sub-suppliers. Such mechanisms may include self-assessments, observations, testing and external audits.

References

Any tenderer bidding on the development and delivery of an AI-based system or solution should demonstrate a sufficient level of experience with the application of AI/ML in a similar or comparable sector or thematic area. That level of experience should be demonstrated by suitable references from past contracts.

General requirements for public procurement, such as the establishment of any conflicts of interest, may be supplemented with considerations regarding ethical impacts or risks linked to fulfilled contracts if they are part of thorough, independent evaluations of projects or external audits.

RECOMMENDATION: Procurement process - selection criteria

Selection criteria	Content
A diverse, multidisciplinary team	<ul style="list-style-type: none"> ▪ AI system and data engineering ▪ Data science ▪ Coding ▪ Testing ▪ Law ▪ IT and cyber security ▪ Change management ▪ Monitoring mechanisms ▪ Process facilitation ▪ Training and awareness raising
Technical facilities	<ul style="list-style-type: none"> ▪ Development environment ▪ Test environment ▪ Physical and IT security measures ▪ Location within or outside the EU/EEA
Sub-suppliers	<ul style="list-style-type: none"> ▪ Location within or outside the EU/EEA ▪ Risk assessments ▪ Control mechanisms, including audits
References	<ul style="list-style-type: none"> ▪ Past AI contracts ▪ Independent evaluations and audits of projects

8.3. Technical specifications

The technical specifications of all tenders for AI-based services or solutions should include requirements regarding the methodologies and processes planned for the development of that AI-based system or solution. This should address all aspects of the contract and cover any stage during the design, development, training and deployment of the service or system.

In that light, public procurement processes should be in place to consider and assess if the tenderer of AI-based services and solutions meets the following supplementary requirements.

Applied standards

If standardized management systems are applied to quality, information security, cyber security, cloud solutions, privacy/PII, data ethics, the environment, social responsibility and AI, the tenderer should identify and describe applied standards, e.g. ISO standards, OECD guidelines, IEEE recommended practices or their equivalents.

The tenderer should demonstrate that certifications or official recognition/approvals have been obtained or granted from international, regional or national standardization or accreditation institutions or their equivalents.

Fundamental rights compliance

Due to the inherent high risk that AI-based systems and solutions pose in causing direct or indirect discrimination, unequal treatment, inequality, exclusion and stigmatization, the tenderer should - as a minimum requirement - demonstrate compliance with EU anti-discrimination legislation.

If the AI-based services or solutions entail negative impact on other fundamental rights, especially the right to privacy, free movement and freedom of expression, the tenderer must demonstrate compliance with such rights.

Also, the tenderer should demonstrate how, once identified, risks to fundamental rights will be handled, i.e. eliminated, minimized or prevented in the design phase and when developing, training and deploying the AI-based service or solution, and through which methodologies, tools and measures (e.g. by design strategies and design patterns that ensure privacy and data protection).

Universal design

The tenderer should explain and demonstrate how accessibility for persons with disabilities or how 'design for all' are integrated in the design, development and testing of the proposed AI-based service or solution. This includes explaining the choices and outcomes of the service/solution that could impact its accessibility, and any potential mitigating measures.

Data ethics management

Equivalent to the requirement to demonstrate applied standards, a requirement to demonstrate data ethics management should form part of the technical specifications, as it may reveal that the tenderer has (or lacks) the technical capability to actually fulfil the contract. Such a management system should encompass a data ethics policy, the governance structure, and a description of the roles and responsibilities for each function or task covered by the policy, due diligence processes, monitoring and complaint mechanisms.

Environmental management

In public procurement contracts for AI-based services and solutions, the technical specs should include a requirement regarding the application of environmental management measures or frameworks.

Thus, the tenderer should demonstrate it has measures or structures in place that integrate procedures and processes for staff training, and monitoring, summarizing, and reporting on specialized environmental information relevant to the AI-based service/solution. The application of environmental criteria could encompass, for example, the Ecolabel, which is relevant in relation to hardware, office/lab furniture and equipment, the carbon footprint of the AI development project and the resulting solution's maintenance.

RECOMMENDATION: Procurement process - technical specifications

Technical specifications	Content
Applied standards	<ul style="list-style-type: none"> ▪ ISO standards ▪ OECD ▪ IEEE ▪ EU HLEG on AI ▪ National or regional sector-specific guidelines
Fundamental rights compliance	<p>Demonstration of compliance:</p> <ul style="list-style-type: none"> ▪ Anti-discrimination legislation ▪ Gender equality ▪ Privacy and data protection ▪ Freedom of movement ▪ Risk handling
Universal design	<ul style="list-style-type: none"> ▪ Demonstration of means to ensure accessibility for disabled persons
Data ethics management	<p>Demonstration of:</p> <ul style="list-style-type: none"> ▪ A data ethics policy ▪ Governance structure ▪ Monitoring mechanisms ▪ A complaint mechanism
Environmental management	<p>Demonstration of:</p> <ul style="list-style-type: none"> ▪ Environmental measures or guidelines ▪ Environmental criteria

8.4. Award criteria

Tenders should be assessed according to a set of economic and quality criteria set up by the contracting authorities. They should be transparent and ensure effective, fair competition among the tenderers.

The contracting authorities should also ensure that it is possible to effectively verify the information provided by tenderers and facilitate the process of identifying the most economically advantageous tender, i.e. the tender that demonstrates the best price-performance ratio.

By awarding the contract, the contracting authorities should ensure that objective criteria are applied in compliance with the principles of transparency, non-discrimination and equal treatment, and that they are aimed towards an overall goal of sustainability.

Quality

The quality criteria of a request for tenders should reflect the technical specifications regarding applied standards and management systems for categories such as information security, data ethics, the environment, privacy and universal design.

Quality criteria could also include the ability of the AI model to adapt to change, e.g. through monitoring the model management and key figures.

Also, they could include the tenderer's technical merits within AI/ML, and the aesthetic, functional and innovative characteristics of the proposed service or solution.

Other quality criteria relevant to the subject matter of the contract could include the handling or integration of data ethics, environmental and social aspects, as revealed in the impact assessments performed as part of the due diligence process (see section 5).

It may be necessary to operate with a non-exhaustive list of potential quality criteria to ensure a comparative assessment of the performance offered by each tender.

As with general awarding processes in public procurement, the assessment of the best price-performance ratio should define the economic and qualitative criteria linked to the subject matter of the contract and reflect the specific characteristics of the AI-based service or solution.

Organization

Designing, developing, testing and deploying AI-based services and solutions typically requires staff members who possess specific qualifications and experience. In public procurement contracts in particular, the tenderers' merits and experience within accessibility and universal design, data ethics, environmental and social impact assessments, fundamental rights compliance analyses and risk assessments, including privacy, data protection, gender equality, non-discrimination and equal treatment, are pivotal to the quality of the outputs and outcomes generated by the service or system.

Public-sector AI-based services and solutions to automatize or support decision making in cases or processes involving citizens will affect the lives of men, women and those with other gender identities. To avoid discriminatory gender bias in designing such services or solutions and choosing and coding the variables in the algorithm, it may be relevant to add the demonstration of gender equality in design and development teams to the list of award criteria.

Similarly, any potential adverse impact on citizens due to their age, race, ethnicity, religion or belief, disability, or sexual orientation should be eliminated or mitigated when developing AI-based services and solutions through the tenderers' demonstration of diversity in all DevOps teams.

Best price-performance ratio

In order to assess the best price-performance ratio, the applied qualitative criteria should be accompanied by a cost criterion, e.g. the price or cost defined on the basis of a cost-effectiveness approach (such as life-cycle cost analysis).

RECOMMENDATION: Procurement process - award criteria

Award criteria	Content
Quality	<ul style="list-style-type: none">▪ Technical merits within AI/ML▪ Experience with model management systems▪ Functionality, aesthetics and innovative characteristics▪ Integration of data ethics, environmental and social aspects
Organization	<ul style="list-style-type: none">▪ Data ethics, environmental and social impact assessments skills▪ Fundamental rights compliance and risk assessment skills▪ Non-discrimination and equal treatment skills▪ Gender balance in teams▪ Diversity in teams
Best price-performance ratio	<ul style="list-style-type: none">▪ Price or cost of each qualitative criterion

9. Contract Performance Conditions

To fulfil the overall goal of sustainability and respect for fundamental rights, and the specific goal of data ethics in AI-based services and solutions, the contracting authority should include clauses in the contract performance conditions on these issues, together with possible penalties and documentation requirements.

Contract clause

The data ethics contract clause should include a requirement regarding relevant management systems or structures with the contracting party to ensure - as a minimum - data ethics, legal compliance, accountability, technical robustness and sustainability.

Specific requirements regarding relevant information security standards and specific legal frameworks may be added.

As part of fulfilling such requirements, the contract party should be under an obligation to ensure appropriate technical and organizational measures with the contracting party.

For example:

- Due diligence processes
- Compliance analyses
- Data ethics policies and procedures
- Data protection
- Information and cybersecurity
- Procedures for integrating data ethics principles and data protection requirements throughout the project's life cycle, i.e. in the design phase and during development, testing, launching, deployment and maintenance
- Any relevant requirements regarding the team's diversity and gender equality could be included.

If the supplier uses sub-suppliers to carry out specific tasks under the public contract, an obligation should be imposed on the supplier to provide sufficient guarantees to establish appropriate technical and organizational measures with the sub-suppliers that correspond to those imposed on the supplier.

Penalties and termination

Contract performance conditions should include penalties for non-compliance. They should not be imposed immediately but seen as a process initiated by the contracting authority with the purpose of finding constructive solutions to problems and challenges in relation to data ethics, compliance, accountability, technical robustness and sustainability.

Such a process should be based on dialogue and engagement with the goal of finding agreed-upon solutions. It may include a wide range of dialogue and mediation tools and approaches and may also include other stakeholders. The process could cover initiatives that mitigate the negative impact on fundamental rights, preventive mechanisms, those that improve the integration of data ethics in the system's design, other forms of reparations, or - as a last resort - termination of the contract and a claim for damages.

Documentation

The supplier should demonstrate accountability across all phases and all levels of the AI-based service/solution and its development. Documentation requirements should encompass:

- Governance structure
- Management systems or frameworks
- Applied standards
- Methodologies
- Analyses and assessments
- Procedures and processes
- Risk-management initiatives
- Monitoring/auditing mechanisms
- Training and testing
- Evaluations.

Other activities or incidents related to contractual requirements on data ethics, compliance, accountability, technical robustness and sustainability should also form part of the contract.

The public authority should have access to documentation upon request and no later than 30 days after a written notification is delivered to the supplier.

Inspection and monitoring mechanisms

The contract should include a clause regarding monitoring and auditing mechanisms in relation to appropriate fulfilment and due diligence of contractual requirements relating to data ethics, compliance, accountability, technical robustness and sustainability.

Inspections should be carried out on a regular and systematic basis and may be based on a self-assessment by the supplier, on-site observation by the commissioning public authority, or by an external and independent evaluation or audit.

The contract may also specify monitoring mechanisms to be applied to sub-suppliers.

RECOMMENDATION: Procurement process - contract performance conditions

Contract performance conditions	Content
Contract clauses	<ul style="list-style-type: none"> ▪ Management systems ▪ Technical and organization measures ▪ Data ethics policy ▪ Sup-suppliers
Penalties	<ul style="list-style-type: none"> ▪ Dialogue and engagement ▪ Mitigation of negative impact ▪ Preventive mechanisms ▪ Termination ▪ Damages
Documentation	<ul style="list-style-type: none"> ▪ Life cycle documentation ▪ Access to documentation
Inspection and monitoring mechanisms	<ul style="list-style-type: none"> ▪ Regular and systematic inspections ▪ Verification mechanism regarding sub-suppliers

10. Contract Implementation

When implementing contracts for AI-based services and solutions, it is necessary to focus on the thematic areas covered by data ethics in public procurement as laid out in this white paper.

To meet the requirements established by a public contract under the categories of data ethics, legal compliance, accountability, technical robustness and sustainability, the supplier will have to set up an organization to monitor and provide insight into its contractual obligations and corresponding processes.

Similarly, the contracting public authority should be engaged in and have access to processes and tools that enables it to identify, evaluate and report on progress, documentation and potential risks with the supplier for not fulfilling the contract.

Organizational measures

The effective implementation of a public contract with a supplier presupposes clearly-defined work processes within a managerial structure, such as a project management unit or a programme management office, depending on the size of the project.

A governance structure that includes top management in decision-making processes should support the AI project and identify roles and responsibilities in relation to all its phases and levels.

The resulting organization should be designed to oversee development processes, administrative processes, evaluations and inspections, and reporting schemes internally and with the public authority. It may also involve training employees in areas covered by the contract, e.g. impact assessments on data ethics and fundamental rights.

Apart from providing support to all employees and teams within the project, these organizational measures should ensure that the contract's requirements are fulfilled, and that anomalies, delays or other relevant impediments are reported to the public authority.

Depending on the project, the public authority may play an important part in the development and deployment process, e.g. by participating in steering committees or acting as the project manager/owner. In that case, all organizational measures should reflect the collaboration or partnership between the supplier and the public authority.

Follow-up mechanisms

The public authority procuring an AI-based service or solution should have follow-up mechanisms in place to effectively monitor and check on progress and results, and the fulfilment of the contractual requirements.

Follow-up measures may include initiation meetings, regular status meetings and continuous dialogue with the supplier. Certain parts of the initial phases of the AI project may require the inclusion of stakeholders, e.g. groups of citizens, NGOs, consumer organizations, academia, and other public authorities (data protection authorities, national human rights institutions, etc.) to elaborate on potential adverse effects on minority groups.

Also, the contracting public authority may have an interest in testing and evaluating applied methodologies, model management, coding reviews, or may have insight into training data etc., or into the documentation of certification updates and other matters.

Monitoring mechanisms may also include self-assessments performed by the supplier in relation to some or all phases of the project. Inspections could also be performed by third parties or as part of external and independent audits.

As part of contract management, the contracting public authority must often approve the use of new sub-suppliers. This requires formalized processes on both sides.

RECOMMENDATION: Procurement process - contract implementation

Implementation of contracts	Content
Organizational measures	<ul style="list-style-type: none"> ▪ Formal procedures ▪ Defined work processes ▪ Monitoring mechanisms ▪ Reporting structure ▪ Contract management system ▪ Training
Follow-up	<ul style="list-style-type: none"> ▪ Start-up and status meetings ▪ Self-evaluation ▪ Continuous dialogue ▪ Model management ▪ Certification updates ▪ Evaluation and audits ▪ Approval of sub-suppliers

11. Recommendations

Inclusion of data ethics in the procurement process

This white paper recommends the inclusion of data ethics requirements in the public procurement of AI-based services and solutions.

To this end, a procurement model should include a **due diligence process**. This makes it possible for the public procurer to gain qualified insight into the planned AI model, its risks and impacts. It would also provide the public procurer with important information about potential suppliers in the market, their skills, methodologies and management systems.

Due diligence helps ensure that all relevant risks and impacts relating to data ethics are identified and handled, and that the development, deployment and maintenance of AI accommodates the requirements of a trustworthy AI-based system or solution. As such, a properly diligent procurement process spans the whole lifecycle of the AI-based system or solution.

Recommendations are made in relation to five phases of the procurement process:

1. Preliminary risk and impact assessments
2. Preliminary screening of potential suppliers
3. Contracting
4. Contract performance conditions
5. Contract implementation.

All recommendations address how to include data ethics in the development and operations of AI-based services and solutions, identifying the relevant processes, tools and methodologies to integrate data ethics requirements and demonstrate their fulfilment.

The need for a common standard

This whitepaper demonstrates the need for a common **standard for data ethics within the EU for the public procurement of AI-based services and solutions**. Such a step would further a human-centric, sustainable usage of AI in the public sector and provide leverage to a business community that is eager to sell AI-based services and solutions but lacks insight into fundamental rights, IT security and data ethics.

This whitepaper suggests that the following initiatives are integrated and promoted within the European Union's Public Procurement Strategy.

These recommendations encompass the integration of data ethics principles in regulation, political priorities, training and awareness-raising initiatives.

- **An EU directive on the public procurement of AI-based services and solutions in the public sector**

A new directive or amendments to existing EU regulations should legally require member states to consider data ethics when procuring trustworthy AI for public sector purposes. The requirement could be integrated into all phases of the procurement process. Such integration should be guided by a due diligence approach, as illustrated by this white paper.

- **Guidelines on the public procurement of AI-based services and solutions in the public sector**

As an alternative or supplement to an EU directive, official EU guidelines should be adopted. The main purpose should be to guide the establishment of procurement and due diligence processes in public sector. It should illustrate how the inclusion of data ethics in the aforementioned processes may lead to procurement of trustworthy AI, including data ethics.

Also, these guidelines should inform private sector suppliers about requirements and expectations regarding their governance and management schemes in relation to design, development, deployment and maintenance of AI-based services and solutions.

- **Inclusion of data ethics as a strategic policy priority in the Public Procurement Strategy of the EU**

The EU Procurement strategy reflects the need to consider the environmental and social impact of new products and services as an integrated part of the procurement process. In light of the rapid development of AI and the expected influence on the decision-making processes of public administrations, it seems both timely and necessary to include data ethics as a new strategic priority when it comes to policy.

- **A training toolkit on data ethics in the procurement of trustworthy AI-based services and solutions**

To promote the principles of data ethics and improve knowledge and awareness of data ethics and due diligence in public procurement processes, a training toolkit should be developed.

It should address the public sector as the contracting party in the procurement process and the private sector in its role as the bidder on tenders issued by public entities.

- **A data ethics handbook for the procurement of AI-based services and solutions**

Complementary to the training toolkit, a handbook should be distributed. It should explain data ethics principles and how to integrate them when designing, developing, deploying and maintaining trustworthy AI-based services and solutions.

- **An online help desk**

If implemented as suggested, contracting public authorities and private sector tenderers may need help understanding procedures, compliance requests, technical specifications and organizational measures such as processes and the demonstration of the fulfilment of requirements. An online help desk would meet that need and be instrumental for greater, more active engagement in data ethics and trustworthy AI in the EU.

Sources

Reports & frameworks:

AI Auditing Framework, ICO, 2019

(including <https://ico.org.uk/about-the-ico/news-and-events/ai-blog-an-overview-of-the-auditing-framework-for-artificial-intelligence-and-its-core-components/>)

Opinion of the Data Ethics Commission, 2019

https://datenethikkommission.de/wp-content/uploads/191023_DEK_Kurzfassung_en_bf.pdf

Automating Society: Taking Stock of Automated Decision-Making in the EU, AlgorithmWatch in cooperation with Bertelsmann Stiftung, 2019.

Data Flows - Future Scenarios: In-Depth Analysis for the ITRE Committee, Simon Forge & Collin Blackman, 2017.

Communication: A European strategy for data, European Commission, 2020.

Ethics Guidelines for Trustworthy AI, European Commission's High-Level Group on AI, 2019.

Menneskerettigheder og offentlige indkøb, Institut for Menneskerettigheder, 2019

OECD Guidelines for Multinational Enterprises, OECD 2011

OECD Due Diligence Guidance for Responsible Business Conduct, 2018

Policy and investment recommendations for trustworthy Artificial Intelligence, 26 June 2019. European Commission's High-Level Group on AI.

UN Guiding Principles for Business and Human Rights, 2011

White paper on artificial intelligence - A European approach to excellence and trust, European Commission, 2020

World Economic Forum, Guidelines for AI Procurement - Whitepaper, September 2019

Books:

Håndbog i dataansvarlighed, Djøf 2020. Birgitte Kofod Olsen.

Data Ethics - The New Competitive Advantage, 2016. Gry Hasselbalch & Pernille Tranberg.

DATAETHICS

[Dataethics.eu](https://dataethics.eu)

info@dataethics.eu

CVR 38465724

ISBN 978-87-972168-0-4