ELI
EUROPEAN
LAW
INSTITUTE

Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration

Report of the European Law Institute

ELI

EUROPEAN
LAW
INSTITUTE

Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration

Report of the European Law Institute

The European Law Institute

The European Law Institute (ELI) is an independent non-profit organisation established to initiate, conduct and facilitate research, make recommendations and provide practical guidance in the field of European legal development. Building on the wealth of diverse legal traditions, its mission is the quest for better law-making in Europe and the enhancement of European legal integration. By its endeavours, ELI seeks to contribute to the formation of a more vigorous European legal community, integrating the achievements of the various legal cultures, endorsing the value of comparative knowledge, and taking a genuinely pan-European perspective. As such, its work covers all branches of the law: substantive and procedural; private and public.

ELI is committed to the principles of comprehensiveness and collaborative working, thus striving to bridge the oft-perceived gap between the different legal cultures, between public and private law, as well as between scholarship and practice. To further that commitment it seeks to involve a diverse range of personalities, reflecting the richness of the legal traditions, legal disciplines and vocational frameworks found throughout Europe. ELI is also open to the use of different methodological approaches and to canvassing insights and perspectives from as wide an audience as possible of those who share its vision.

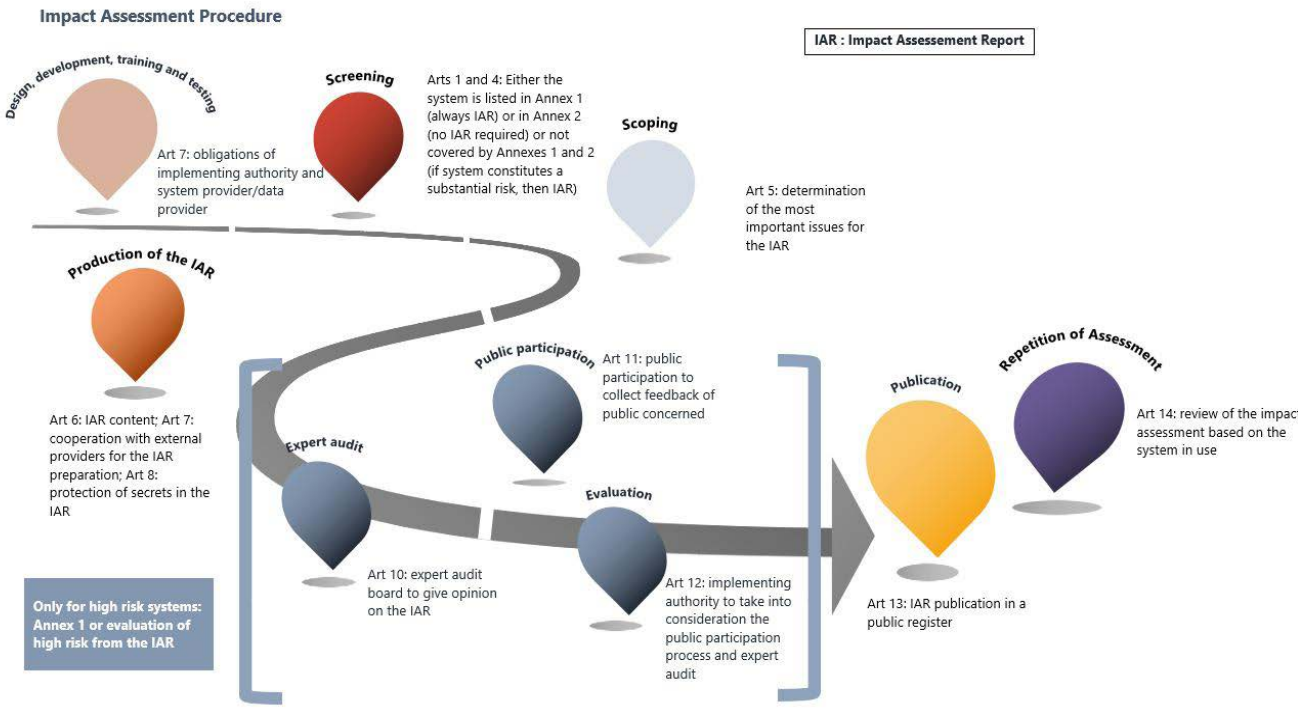
European Law Institute Secretariat
Schottenring 16/175
1010 Vienna
Austria
Tel.: + 43 1 4277 22101
Mail: secretariat@europeanlawinstitute.eu
Website: www.europeanlawinstitute.eu

ISBN:978-3-9505192-1-1

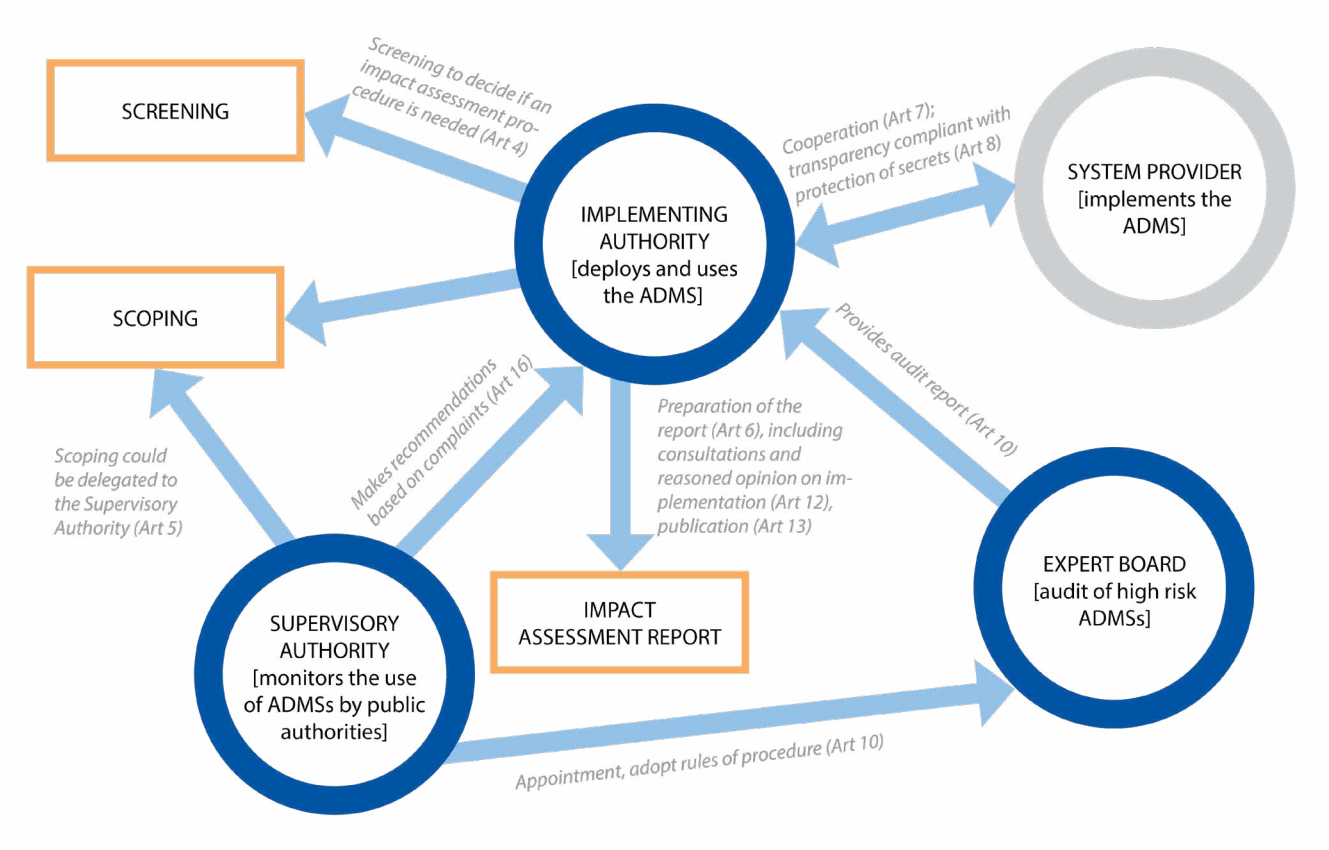
Acknowledgment is due to the University of Vienna, which has generously hosted the ELI Secretariat under successive Framework Cooperation Agreements since 2011.

universität wien

Table of Contents



Timeline of development of the Impact Assessment Procedure



Role of bodies involved in the Impact Assessment Procedure

Acknowledgements

Project Team
Chair
Marek Wierzbowski (Lawyer and Professor, Poland)

Project Reporters
Marc Clément (Administrative Court Judge, France)
Paul Craig (Professor, United Kingdom)
Jens-Peter Schneider (Professor, Germany)

Other Members
Jonathan Dollinger (Research Assistant and Doctoral Candidate, Germany; 'Person With the File' (PWF))
Franz Merli (Professor, Austria)
Daniel Le Métayer (Senior Research Scientist, France)
Karolina Wojciechowska (Researcher and Assistant Professor, Attorney-at-Law, Poland)
Katarzyna Ziółkowska (Doctoral Candidate, Poland)

Advisory Committee
Assessors
Philip Moser QC (Barrister, United Kingdom)
Pascal Pichonnaz (Professor, Switzerland)
Ilaria Pretelli (Legal Adviser, Switzerland)
Fryderyk Zoll (Professor, Germany/Poland)

Other Members
Fabrizio Cafaggi (Member of the Italian Council of State, Italy)
Raja Chatila (Professor, France)
Péter Darák (Supreme Court Judge, Hungary)
Jonas Ebbesson (Professor, Sweden)
Joanna Goodey (Head of Unit, Research and Data Unit, European Union Agency for Fundamental Rights (FRA))
Michael Gøtze (Professor, Denmark)
Alexia Maniaki-Griva (Head of the Ex-Ante Impact Assessment Unit, Directorate-General for Parliamentary Research Services (DG EPRS))
William M McKechnie (Supreme Court Judge, Ireland)
David Reichel (Project Manager, Social Research, Research and Data Unit, FRA)
Jane Reichel (Professor, Sweden)
Jason M Schultz (Professor, Unites States)
Yannick Meneceur (Policy Adviser on Digital Transformation and Artificial Intelligence, Council of Europe)
Olivia Tambou (Associate Professor, France)
Clara Velasco (Associate Professor, Spain)

Members Consultative Committee
Raquel Abajas (Legal Assistant, Spain)
Chiara Silvia Armida Angiolini (Independent Researcher, Italy)
Francesco Avolio (Lawyer, Italy)
Arvind Babajee (Consultant, Mauritius)
Alexander Balthasar (Visiting Professor, Austria)
Robert Bray (Lawyer and Linguist, Belgium)
Alessandro Cenerelli (Doctoral Candidate, Italy)
Ignace Claeys (Professor, Belgium)
Council of the Notariats of the European Union (represented by Tamás Parti, Raul Radoi and Tamás Sajben)

Edita Čulinović Herc (Professor, Croatia)
Alberto De Franceschi (Professor, Italy)
Gudula Deipenbrock (Professor, Germany)
Darinka Dekleva Marguč (Senior Administrative High Court Judge, Slovenia)
Mustafa Ebaid (Legal Researcher, Turkey)
Wian Erlank (Professor, South Africa)
Dessislava Fessenko (Attorney, Bulgaria)
Elena Mihaela Fodor (Associate Professor, Romania)
Laurence Gormley (Professor, The Netherlands)
Patrícia Guimarães (Associate Professor, Brazil)
Sarah Houllier (Administrative Court Judge, France)
Dariusz Kloza (Postdoctoral Researcher, Belgium)
Maria Lubomira Kubica (Assistant Professor, Spain)
Stephanie Laulhe Shaelou (Professor, Cyprus)
Dulce Lopes (Assistant Professor, Portugal)
Elwira Macierzyńska-Franaszczyk (Assistant Professor, Poland)
Caroline Mantl (Senior Legal Expert, Ireland)
Carlos Marinho (Court of Appeal Judge, Portugal)
Irena Nesterova (Postdoctoral Researcher, Latvia)
Elena Alina Ontanu (Assistant Professor, The Netherlands)
Manuel Peláez Muras (Senior Expert in Public Procurement, Spain)
Katarzyna Pokryszka (Lecturer, Poalnd)
Radim Polčák (Professor, Czech Republic)
Oreste Pollicino (Associate Professor, Italy)
Francesco Quarta (Assistant Professor, Italy)
Teresa Rodriguez de las Heras Ballell (Associate Professor, Spain)
Leigh Sagar (Barrister, United Kingdom)
David Michael Schneeberger (Research and Teaching Assistant, Austria)
Anna Simonati (Professor, Italy)
Guillem Soler Solé (Judge, Spain)
Henrique Sousa Antunes (Professor, Portugal)
Sjef van Erp (Professor, The Netherlands)
Veronica Williams (Legal Officer, Belgium)
Boštjan Zalar (High Court Judge, Slovenia)

ELI Project Officer
Katja Kolman (Senior Project Officer, Austria)

During the course of the project the Project Team was kindly assisted by the ELI Bodies and the ELI Secretariat in its work.

Participants in Project Team Meetings
1. 15 April 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
2. 28 April 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
3. 26 May 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
4. 1 July 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
5. 25 August 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
6. 8 October 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.

7. 16 November 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
8. 9 December 2020: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
9. 26 January 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Olivia Tambou, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
10. 25 February 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
11. 29 March 2021: Marc Clément, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
12. 23 April 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Daniel Le Metayer, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
13. 25 May 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
14. 21 July 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
15. 18 August 2021: Marc Clément, Paul Craig, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
16. 30 September 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.
17. 24 November 2021: Marc Clément, Paul Craig, Jonathan Dollinger, Franz Merli, Jens-Peter Schneider, Marek Wierzbowski, Karolina Wojciechowska, Katarzyna Ziółkowska.

27 October 2020: Meeting with the Advisory Committee
29 June 2021: Meeting with the Advisory Committee
27 October 2021: Meeting with the Advisory Committee

24 November 2020: Meeting with the Members Consultative Committee
3 November 2021: Meeting with the Members Consultative Committee

The views set out in this report should not be taken as representing the views of those bodies on whose behalf individual members of the Project Team and Advisory Committee were also acting.

Executive Summary

Public bodies have made decisions and rules from time immemorial. The nature of the particular public bodies perforce varies as between legal systems. It is, in addition, commonplace to accept that decisions may also be subject to public law in certain instances when they have been made by private actors. There is nonetheless a foundational commonality underlying the preceding heterogeneity, which is that while the institutions might have varied in certain respects as between legal systems, the decisions were made by human beings. There was an individual, or institution, that made the contested rule or decision. The subject matter of these Model Rules attests to an important change in this regard, since the reality is that in many instances it is not possible to trace a decision back to a discrete individual. The operative decision may be made by an algorithm, or some other form of automated decision-making. Humans may still be involved in such decisional processes, in the sense that they may design the algorithm, and there may also be some human involvement before the operative decision is taken. However, a system may be fully automated when set up, such that the output/decision can occur without human involvement, and some systems make provision for the algorithm to learn and develop. It is, therefore, unsurprising that the existence of such automated systems broadly conceived poses novel problems for both public and private law. This is attested to by the plethora of initiatives dealing with such matters emanating from bodies such as the European Union (EU) and the Council of Europe, as well as from particular nation States.

It is important to underline that these Model Rules have been designed so that they are not dependent on EU law and can be implemented in non-EU legal systems. In other words, they must be able to fit into different legal contexts that do not incorporate fundamental elements of EU data protection law such as the General Data Protection Regulation (GDPR). However, the rules presented have been developed in such a way as to ensure that they are compatible not only with existing EU law, but also with the law currently being drafted, in particular the Draft Regulation on Artificial Intelligence (AI). The latter draft concerns to a large extent the AI projects developed or used by administrations. The Model Rules complement the Draft Regulation's approach by providing specific safeguards for democracy, the right to good administration and the rule of law when algorithmic decision-making systems are used by the public administration. This is exemplified by the provision made for case-specific impact assessments including public and expert participation.

There are various ways in which concerns raised by algorithmic decision-making can be addressed. The central idea underlying these Model Rules is for an Impact Assessment to be conducted. The very variety of situations in which algorithmic decision-making is employed precludes a one-size fits all approach. This would lead to rules that were too rigorous for some such systems, and too generous for others. The approach in the Model Rules is therefore variegated. A legal system that adopts the Model Rules can specify that certain such systems fall into Annex 1, which denotes that they are high risk, and hence always subject to an Impact Assessment. It can, to like effect, stipulate that other systems should fall into Annex 2, because they are regarded as low risk and therefore do not warrant such an Assessment. There may, however, be other such systems that cannot readily be classified ex ante as falling within either Annex 1 or Annex 2. These systems are subject to an initial risk evaluation in accordance with a screening procedure. An Impact Assessment is then required if the system constitutes at least a substantial risk according to the screening procedure.

The Model Rules set out in detail the nature of the Impact Assessment, which is framed by a prior scoping procedure designed to target the more particular issues on which such an Assessment should focus. The Impact Assessment is intended to be both measured, in the sense of addressing the benefits as well as the risks from the use of algorithmic decision-making; and proportionate, in the sense of not being too burdensome for the authorities that have to conduct them. There are, however, further requirements for such systems that are regarded as high risk, either because they fall within Annex 1, or because they are deemed to be so as a result of a particular Impact Assessment. The further requirements are set out in Chapter 3 of the Model Rules and entail, inter alia, scrutiny of the Impact Assessment by an expert board, as well as the opportunity for public participation. The Model Rules also make provision for a Supervisory Authority to oversee the preceding processes, and specify the circumstances in which a legal challenge might be made.

Reporters' Preface

Public administration, being an emanation of the State's public functions, entails the processing of much more data than most private entities. New technologies, such as AI, can therefore play a significant role in the modernisation and overall improvement of the functioning of public administration. On the other hand, a guarantee of the transparency, correctness and security of the processed data is also fundamental. Therefore, the possibility to implement AI in the operation of public administration is limited by the principle of legality and the need to ensure a high degree of reliability of technologies used, as well as the need to ensure respect for citizens' rights.

Public administration is, as a result, confronted with specific challenges in the deployment of AI and, more generally, algorithmic decision-making systems (ADMSs), even if they do not use specific AI technologies, such as machine learning. The use of these techniques poses specific problems related to the principle of good administration. In addition, issues such as transparency, accountability, compliance and non-discrimination are particularly relevant in the context of public administration. These Model Rules aim to lay down the foundation for supplementing European legislation on AI in the specific context of public administration that will not hinder innovation, while providing solid safeguards to improve citizens' confidence in the use of the technology in this field, by promoting the role of impact assessment.

Impact assessment is, therefore, a tool to analyse the effects of ADMSs used by public authorities.

The impact assessment should:

- raise awareness of the risks of ADMSs in public administration;
- enable the administrative authorities to make an informed decision on the use of ADMSs;
- allow experts and the public to participate in the decision-making process;
- render the decision-making process and its result more transparent for the public; and
- make it easier to hold the public administration accountable for the use of ADMSs.

Impact assessment is not a licensing procedure. It results in a report and not a licence. Compliance with legal requirements is one assessment criteria for a proposed use of an ADMS, but the assessment does not produce a binding decision on its legality. An impact assessment is mandatory for certain ADMSs, so their use without a preceding assessment would be illegal. However, the assessment does not legally determine the public authority's decision to use a certain ADMS, but leaves this decision to the authority's discretion.

Although the Model Rules have been developed with some inspiration from EU law, the Model Rules are intended to be more general and adaptable in different legal contexts within and beyond the EU.

An important question is which entities could adopt the Model Rules. Several options exist for the EU and its Member States. The EU clearly has the legislative power to mandate their institutions, bodies and agencies to conduct an ADMS impact assessment in line with the Model Rules. A more complex issue is the legislative competence of the EU concerning an obligation for national authorities to conduct an ADMS impact assessment. This important legal question needs further consideration and goes beyond the scope of this ELI project. The legal status quo is that the EU Member States are responsible and competent to oblige their public authorities to conduct an ADMS impact assessment. On the national level, whether the Model Rules are implemented in national law, by a regional sub-division or on both levels depends on the division of legislative and administrative competences.

According to Article 1 (a) of its current Draft AI Regulation, the Commission proposes to regulate not just the development and marketing, but also the use of AI. If the EU follows this approach, the Project Team would recommend the inclusion of opening clauses that allow Member States to enact ADMS impact assessment rules following the Model Rules. This would not be an undue burden on the common market: An impact assessment does not concern the production or marketing of AI. It also does not create substantive rules, unlike Title III, Chapter 2 of the Draft AI Regulation.

However, if the result of an in-depth analysis of the EU Treaties is that the EU is competent to oblige Member States to conduct impact assessments for ADMSs used by their national authorities, the ELI Project Team would be in favour of an EU-wide implementation of the Model Rules. In this case, the Draft AI Act might allow the Member States to specify details, especially about the applicability and scope of the impact assessment by modification to Annexes I and II. This would empower them to take into account their respective societal, cultural and technological context.

List of Sources

EU Legislation

Council Directive 91/250/EEC of 14 May 1991 on the legal protection of computer programs [1991] OJ L122/42.

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20.

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) [2007] OJ L108/1.

Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs [2019] OJ L111/16.

Directive 2011/92/EU of the European Parliament and of the Council of 13 December 2011 on the assessment of the effects of certain public and private projects on the environment [2011] OJ L26/1.

Regulation (EU) 182/2011 of the European Parliament and of the Council of 16 February 2011 laying down the rules and general principles oncerning mechanisms for control by Member States of the Commission's exercise of implementing powers [2011] OJ L55/13.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data 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] OJ L119/89.

Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union [2018] OJ L303/59.

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information [2019] OJ L172/56.

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM (2021) 206 final.

Other Legislation

Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters (Aarhus Convention).

Directive on Automated Decision-Making (Canada), <https://tbs-sct.gc.ca/pol/doc-fra.aspx?id=32592> accessed 1 October 2021.

Bill Requiring Companies To Target Bias In Corporate Algorithms (USA, Senators Booker and Wyden), < www.booker.senate.gov/?p=press_release&id=903> accessed 1 October 2021.

Senate Bill 5116 (Washington, Senators Hasegawa, Hunt, Kuderer and Wilson), <https://lawfilesext.leg.wa.gov/biennium/2021-22/Pdf/Bills/Senate%20Bills/5116.pdf> accessed 1 October 2021.

Environmental Impact Assessment Act (Gesetz über die Umweltverträglichkeitsprüfung) (Germany).

Reports and Documents

Commission, 'Artificial Intelligence for Europe' COM(2018) 237.

Commission, 'Staff Working Document on Liability for Emerging Digital Technologies' SWD(2018) 137 final.

Commission, 'Building Trust in Human-Centric Artificial Intelligence' COM(2019) 168 final.

Commission, 'White Paper on Artificial Intelligence - A European Approach to Excellence and Trust', White Paper COM(2020) 65 final.

Commission, 'Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics' COM(2020) 64 final.

Independent High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai> accessed 9 September 2021.

Independent High-Level Expert Group on Artificial Intelligence, 'Assessment List for Trustworthy Artificial Intelligence (ALTAI), <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment> accessed 9 September 2021.

EU Fundamental Rights Agency, Getting the future right, <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights> accessed 14 December 2021.

Catelijne Muller, 'The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law' (CAHAI (2020) 06).

Council of Europe, Ad Hoc Committee on Artificial Intelligence, 'Feasability Study' (CAHAI (2020) 23).

Council of Europe, Ad Hoc Committee on Artifical Intelligence, Human Rights, Democracy and Rule of Law, 'Impact Assessment of AI systems' (CAHAI-PDG (2021) 05).

Data Ethics Commission (Germany), 'Opinion of the Data Ethics Commission', <www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html>, accessed 14 September 2021.

Information Commissioner's Office, 'AI Auditing Framework', <https://ico.org.uk/about-the-ico/news-and-events/ai-auditing-framework/> accessed 14 December 2021.

KI Bundesverband, Position Paper on EU-Regulation of Artificial Intelligence, <https://ki-verband.de/wp-content/uploads/2021/02/Final_Regulierung-komprimiert-1.pdf> accessed 14 December 2021.
ReNEUAL, 'Model Rules on EU Administrative Procedure Book II –Administrative Rulemaking' 2014, <http://www.reneual.eu/images/Home/BookII-AdministrativeRulemaking_individualized_final_2014_09_03.pdf> accessed 09 September 2021.

In addition, the Project Team consulted academic literature. However, in order to present the Model Rules in the style of a legislative proposal, the editorial board decided to refrain from references to academic literature.

ELI Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration – Black Letter Rules

Chapter 1: General Provisions

1

Article 1: Purpose and Scope

1. These Model Rules provide for an impact assessment of those algorithmic decision-making systems (hereinafter: system(s)) used by public authorities which are likely to have significant impacts on the public.

2. The implementing authority shall carry out an impact assessment in accordance with Articles 5 to 13 before deploying any system that is listed in Annex 1, or meets the criteria set out in Article 4(1).

3. These Model Rules shall not apply to systems listed in Annex 2.

4. The implementing authority may deploy systems covered by paragraph 2 without a prior impact assessment, or without the consultations mandated in Chapter 3, if

    a. the system is used to respond to an imminent public emergency, in particular relating to public health or safety; and

    b. the system's purpose would be significantly harmed by the delay caused by the impact assessment or the consultation; and

    c. the risk of deploying the system without a prior impact assessment does not outweigh the risk of the delay of the system's deployment.

In this case, the implementing authority shall carry out the impact assessment immediately after it starts using the system.

2

Article 2: Definitions

For the purpose of these Model Rules:

1. 'Algorithmic Decision-Making System' means a computational process, including one derived from machine learning, statistics, or other data processing or artificial intelligence techniques, that makes a decision, or supports human decision-making used by a public authority.

2. 'Public Authority' means:

    a. any government or other public administration, including public advisory bodies, at European Union, national, regional or local level;

    b. any natural or legal person performing public administrative functions under European Union or national law; and

    c. any natural or legal person having public responsibilities or functions, or

providing public services under the control of a body or person falling within (a) or (b).

3. 'Decision' means any determination by a public authority to take or not to take action.
4. 'Public' means one or more natural or legal persons, whether incorporated or unincorporated, and, in accordance with national legislation or practice, their associations, organisations or groups.
5. 'System Provider' means an entity, a department or other organisational unit within an entity, that is responsible for any of the following: designing, developing, setting up, or testing of the algorithmic decision-making system.
6. 'Data Provider' means an entity, a department or other organisational unit within an entity, that supplies data to the system provider, or the implementing authority, for the purpose of designing, developing, setting up, testing, and using the algorithmic decision-making system.
7. 'Implementing Authority' means a public authority that is using, or intends to use, an algorithmic decision-making system. If a superior authority decides to instruct subordinate authorities to use the system, the superior authority is regarded as the implementing authority.
8. 'Supervisory Authority' means the public authority defined in Article 15.

3

Article 3: Coordination with Other Procedures

If an impact assessment under these Model Rules is to be completed, the implementing authority may refer to impact assessments or equivalent documents required by other legislation, such as data protection regulations or product safety regulations. In such cases, the implementing authority shall indicate which part of the Article 6 report shall be substituted by these documents. These documents shall be part of the Article 6 report.

Chapter 2: Standard Impact Assessment Procedure

4

Article 4: Screening

1. Systems not listed in Annex 1 or Annex 2 shall be subject to an initial risk evaluation in accordance with Annex 3 (screening procedure). The implementing authority shall carry out an impact assessment in accordance with Articles 5 to 13 if the system constitutes at least a substantial risk according to the screening procedure.
2. The implementing authority shall publish its answers to the screening questionnaire (screening results) at least online, within two weeks after completion of the screening. Article 13(3) shall apply, mutatis mutandis, to the screening results. The implementing authority shall forward the screening results to the supervisory authority.

5

Article 5: Scoping

1. Before drafting the Article 6 report, the implementing authority may preliminarily identify the most important issues and the necessary level of detail of the impact assessment (scoping). For this purpose, the implementing authority may consult the supervisory authority defined in Article 15, other relevant authorities, the independent expert board established in Article 10, and representatives of the public.

2. The implementing authority may ask the supervisory authority defined in Article 15 to carry out the scoping. In this case, the implementing authority shall take utmost account of the scoping results. If the implementing authority deviates from the scoping results, it shall inform the supervisory authority and give reasons for its deviation. The scoping results shall not bind the supervisory authority in any evaluation of the Article 6 report, or, if applicable, the extended report.

6

Article 6: Impact Assessment Report

1. If an impact assessment is necessary according to Articles 1 or 4, the implementing authority shall prepare an impact assessment report (report).

2. The report shall contain:
   a. a description of the purpose and operation of the system, in particular:
      i. the development of the system, in particular its algorithms;
      ii. the nature and technical characteristics of the system;
      iii. the selection of training, validation and testing data;
      iv. the context in which the system is used, in particular the public objectives as defined in the applicable law;
      v. the system's interrelation with other digital systems deployed by the implementing authority or other public authorities.
   b. an assessment of the performance, effectiveness and efficiency of the system with regard to the public objectives as defined in the applicable law, in particular whether the performance of the system might be flawed by low quality data during its use.
   c. an assessment of the specific and systemic impact of the system on:
      i. fundamental or other individual rights or interests, in particular the rights to privacy and data protection, the right to non-discrimination and the right to good administration;
      ii. democracy, societal and environmental well-being;
      iii. the administrative authority itself, in particular the estimated acceptance of the system and its decisions by the staff, the risks of over- or under-reliance on the system by the staff, the level of digital literacy, and specific technical skills within the authority.
   d. an assessment of the measures taken to ensure:
      i. maximisation of benefits to be achieved by deploying the system with regard to public objectives as defined in the applicable law;

ii. minimisation of identified risks and mitigation of possible negative outcomes;

iii. human agency, oversight and control of the system;

iv. high data quality;

v. accuracy across groups, precision and sensitivity;

vi. technical robustness and safety; resilience to attacks; data security; fall-back plans; reliability; and reproducibility of decisions;

vii. transparency of the system and explainability of its decisions;

viii. traceability to enable the monitoring of the system's operations;

ix. accountability, in particular oversight, auditability, clear allocation of responsibilities, self-monitoring, benchmarking, and the possibility of redress for injury or harm caused by the system;

e. unless the system is listed as 'always high risk' in Annex 1, a concluding determination of the risk level;

f. an overall assessment of the necessity and proportionality of the processing operations in relation to the purposes, in particular trade-offs between different factors set out in this Article and whether there are reasonable alternatives to the envisaged system;

g. a reasoned statement on the legality of the use of the system under the applicable law, in particular data protection law, administrative procedure law and applicable sectoral legislation;

h. any additional information mandated by other Articles of these Model Rules.

3. The assessment guidelines in Annex 4 provide further details on the structure and content of the report. The implementing authority may deviate from these guidelines or use different guidelines if appropriate, especially to account for the particularities of the sector in which the system is used. The implementing authority shall give reasons for any substantial deviation.

4. The report shall describe the system in a manner and depth that is appropriate to the risk level of the system and the context in which it is deployed. The report shall be accurate and understandable to the public. If details are included that are not generally understandable, the implementing authority shall provide a generally understandable summary.

7

Article 7: Cooperation and Communication with the System Provider and the Data Provider

1. The implementing authority and the system provider shall cooperate during the design, development, training, and testing of the system. They shall engage in ongoing communication pursuant to the relevant laws and agreements, in a way that allows adequate reproduction of the information exchange at a later date.

2. The system provider shall provide the implementing authority with all information, documentation, proof, and demonstration necessary for the production of the Article 6 report.

3. The implementing authority and the system provider shall jointly set up a project team and designate their respective representatives to ensure overall, ongoing oversight over the design, development, training, and testing of the system. Representatives designated as members of the project team shall have sufficient expert knowledge to be able to understand the workings of the system, identify desirable results, recognise potential bias, flaws, and malfunctions and report them to the implementing authority. The project team should have unrestricted access to information about the progress of the system development.

4. In order to ensure that the system is working properly under conditions closely reflecting real life operating conditions, the final testing of the system shall be conducted on the implementing authority's equipment, or other equipment used by or on behalf of the implementing authority, by the implementing authority's personnel, with appropriate training by the system provider, including provision of knowledge about the system. The results of the final testing, along with its parameters and testing conditions, shall be included by the implementing authority in the Article 6 report.

5. The implementing authority shall employ suitable clauses in a procurement contract, or take other equivalent action, to ensure the system provider's compliance with the obligations in these Model Rules.

6. Paragraphs 1 to 5 shall also apply, mutatis mutandis, to the data provider.

8

Article 8: Transparency and Protection of Secrets

1. The impact assessment of the system shall be conducted in compliance with the obligation to respect and protect confidentiality of data and information relating or belonging to persons and entities involved in the process of the assessment in accordance with the relevant laws and requirements. This includes, but is not limited to, the protection of personal data, privacy, intellectual property, trade secrets, national security, defence, and public security.

2. The implementing authority and the system provider may reserve confidentiality of data and information shared, used, or acquired throughout the impact assessment other than that referred to in paragraph 1. The reservation of confidentiality of such information and data shall be duly justified and weighed against the requirements of transparency of the impact assessment, the interests of the implementing authority, and of the system provider.

3. The implementing authority shall include in the Article 6 report procedures for accessing the source code of the system and datasets used for system training and testing purposes. The access to the source code, and the training and testing datasets, can be limited or fully restricted when this is necessary to safeguard the legitimate interests and rights of the implementing authority, the system provider or third persons.

4. Paragraphs 1 to 3 shall also apply, mutatis mutandis, to the data provider.

Chapter 3: Additional Provisions for High Risk Systems

9

Article 9: Applicability of this Chapter

The provisions of this Chapter shall apply if:

a. the system is listed as 'always high risk' in Annex 1, or
b. the implementing authority concludes in its Article 6 report that the system constitutes a high risk.

10

Article 10: Expert Audit and Expert Board

1. The Article 6 report shall be audited by an independent expert board.
2. During the expert audit, the expert board shall evaluate the overall quality of the Article 6 report. The criteria for evaluation include accuracy, adequacy, completeness of the impact assessment, and its compliance with these Model Rules. The implementing authority shall provide access to the system for the experts, including the possibility to review the system's source code and datasets used for training and testing purposes, and to use and test the system's working in practice.
3. The results of the expert audit shall be documented in an audit report. The expert board may include in the audit report comments and objections related to, in particular, missing steps of the impact assessment process, deficiencies of the system design, development, training or testing, additional, unforeseen risks, insufficient measures to protect the public, or additional concerns, or recommendations to the implementing authority.
4. Notwithstanding the obligations under Article 12, the implementing authority may respond to the comments and objections raised by the expert board pursuant to paragraph 3 and accordingly complete or amend the Article 6 report. The implementing authority may ask the expert board to give a statement on its response to the audit report.
5. Candidates for the expert board may apply in response to a public call for expression of interest issued by the supervisory authority, or may be proposed by public authorities or the public. The supervisory authority shall hold a list of proposed candidates.
6. The expert board shall be appointed by the supervisory authority from the candidates on the list referred to in paragraph 5 meeting objective criteria specified in advance, on the basis of an open, competitive, non-discriminatory, and transparent procedure.
7. The members of the expert board shall be appointed considering diversity in terms of geography, nationality, vocation, gender, and age, from individuals with relevant knowledge and expertise. They shall have competence and experience, including technological, commercial, business, political, and legal skills, relevant to the use of algorithmic decision-making systems in public administration.
8. In order to avoid any potential conflict of interest and to ensure unbiased audit, the members of the expert board and entities to which they are affiliated cannot

participate in public consultations and cannot be associated, directly or indirectly with the system provider, the data provider and the implementing authority. Where a conflict of interest is deemed to exist, or where one might reasonably be expected to arise, the affected member of the expert board should refrain from participating in the expert audit of the relevant Article 6 report.

9. The supervisory authority shall dismiss a member of the expert board in the event of non-fulfilment or improper fulfilment of the obligations set out in this Article, or in the rules of procedure referred to in paragraph.

10. The supervisory authority shall adopt rules of procedure for the expert board.

11

Article 11: Public Participation

1. After completion of the expert audit, the public shall be consulted by the implementing authority, which shall ensure that those specifically affected by the system are afforded the opportunity to participate in this process.

2. The public and the supervisory authority shall be informed by the implementing authority, at least online, of the following matters expeditiously and, at the latest, when information can reasonably be provided:

   a. the fact that the system is subject to an impact assessment procedure and a short description of the system;

   b. identification and contact details of the implementing authority and details of the time schedule for transmitting comments or questions; and

   c. an indication of the time and place at which, and the means by which, the information as defined in paragraph 3 will be made available.

3. The Article 6 report and the expert audit report (and, where applicable, the response by the implementing authority and the additional expert statement according to Article 10 (4)) shall be made available to the public at least online at the time when the public is informed in accordance with paragraph 2(c). of this Article. If the documents in sentence 1 contain secrets as defined in Article 8, the implementing authority shall make available an edited version of these documents that excludes any secret information.

4. The public shall be given early and effective opportunities to participate in the evaluation of the system and shall, for that purpose, be entitled to express comments to the implementing authority at least by an online consultation exercise.

5. Information and consultation of the public may include a public hearing. There shall be a public announcement of the public hearing, which should be notified at least on an official website.

6. The time-frame for consulting the public shall not be shorter than 30 days after the publication according to paragraph 3.

7. Other public authorities likely to be concerned by the system by reason of their specific responsibilities, including, if relevant, the competent data protection authority, shall also be given an opportunity to express their opinion. The information gathered pursuant to Article 6 shall be forwarded by the implementing authority to those authorities. Paragraphs 1 to 6 shall apply, mutatis mutandis, for this consultation.

12

Article 12: Evaluation and Extended Report

The implementing authority shall consider the information, observations and opinions expressed in the audit report and the public participation process set out in Article 11, and give a reasoned final opinion on the implementation of the proposed system. This evaluation and the final opinion shall be contained in an extended report, which should also include the initial Article 6 report, the audit report and at least a summary of the results of the public participation.

Chapter 4: Conclusion of the Assessment

13

Article 13: Publication

1. The implementing authority shall publish the Article 6 report, or, if applicable, the extended report according to Article 12, at least online and it shall be available for the period when the system is used. After the system is no longer in use, the implementing authority shall retain these documents in accordance with general rules on public files, and at least for one year online.
2. The implementing authority shall give notice of the publication to the experts and members of the public who participated in the consultation under Articles 10 and 11. It shall forward the Article 6 report or, if applicable, the extended report to the supervisory authority.
3. If the documents mentioned in paragraph 1 contain secrets as defined in Article 8, the implementing authority shall publish an edited version of these documents that excludes any secret information. The implementing authority shall grant access to the unedited version of these documents in accordance with the general rules on freedom of information. The supervisory authority shall always receive the unedited as well as the edited version.
4. The documents mentioned in paragraph 1 will be included in the public register under Article 15(4).

14

Article 14: Review and Repetition of the Assessment

1. The implementing authority shall monitor use of the system and review the system whenever there are factual indications of substantial negative impact on aspects covered by Article 6(2)(a) to (d) that were not envisaged in the Article 6 report or, if applicable, the extended report. In particular, it shall review the system if there are changes to the system, the context in which the system is used, or the personnel using it that might have such impact. If the implementing authority finds any such impact, it shall update the Article 6 report or, if applicable, the extended report accordingly.
2. Irrespective of paragraph 1, the implementing authority shall review the system and its impact:

a. after the system has been in use for six months, where it has not been reviewed in accordance with paragraph 1 during the last three months, and

b. every two years after the last review and amend the Article 6 report or, if applicable, the extended report. The amendments shall, in particular, reflect any additional knowledge gained during the practical use of the system.

3. Articles 5, 7 and 8 apply, mutatis mutandis, to such repeat impact assessments. If the implementing authority consulted experts in accordance with Article 10 during the initial impact assessment, it shall forward the amended report to these experts, who may make additional comments. If the implementing authority did not carry out consultations in accordance with Articles 10 and 11 before, but concludes after its review that the system now meets the conditions of Article 9, it shall carry out these consultations.

4. The amended report and, if applicable, any comments made in accordance with paragraph 3, shall be published in accordance with Article 13.

Chapter 5: Accountability

15

Article 15: Supervisory Authority

1. A supervisory authority shall be established to oversee the use of the systems by public authorities.

2. The supervisory authority shall be independent and have adequate financial and human resources to fulfil its tasks under these rules. In particular, it shall be provided with a sufficient number of personnel permanently available, whose competence and expertise shall include an in-depth understanding of artificial intelligence and other algorithmic technologies, data and data computing, fundamental rights, health and safety risks and knowledge of existing standards and legal requirements. The supervisory authority shall have an advisory board composed, among others, of representatives of civil society.

3. The supervisory authority shall:

a. oversee the application of these rules and provisions adopted under these rules;

b. monitor relevant developments concerning the use of a system by public authorities;

c. promote public awareness and understanding of the risks, rules, safeguards and rights in relation to the use of a system by public authorities;

d. upon request, provide information to any affected person on the rules for impact assessment and the use of a system by public authorities;

e. advise public authorities on the impact assessment and the use of a system, in particular on the scoping according to Article 5(2);

f. ensure that sufficient experts are available for audits under Article 10;

g. keep a public register according to paragraph 4;

i. handle, to the extent appropriate, complaints according to Article 16(1);

j. cooperate with other supervisory authorities;

k. issue a yearly report to parliament and the public on its activities and relevant developments concerning the use of a system by public authorities;

l. without prejudice to paragraph 5(e), observe the confidentiality rules set forth in Article 8 in its communications with others.

4. The supervisory authority shall provide a public online register of:

a. screening results according to Article 4(2);

b. ongoing public consultations under Article 11;

c. published reports under Article 13. Article 13(3) shall apply, mutatis mutandis, to documents in the public register.

5. Notwithstanding powers accorded by other Articles in these Rules, the supervisory authority shall have the power:

a. to investigate, on its own initiative or on a complaint under Article 16(1), the application of these Model Rules by the implementing authorities, the appointed experts, and the system or data providers;

b. without limitation by Article 8, to obtain from implementing authorities, appointed experts, and system or data providers all information necessary for the performance of its tasks;

c. to make recommendations to implementing authorities;

d. to initiate, after an unsuccessful recommendation, either legal proceedings before a court of law, or issue a binding order, to stop an implementing authority using a system that, in violation of these rules, has not been subject to a proper impact assessment, or a proper repeat impact assessment;

e. to issue, on its own initiative or on request, opinions to parliament and the government or to other institutions and bodies as well as to the public on any issue related to the use of a system by public authorities.

16

Article 16: Complaints and Legal Protection

1. Without prejudice to more favourable conditions, it shall be ensured that, in accordance with the applicable law, members of the public concerned having a sufficient interest, or alternatively, maintaining the impairment of a right, where administrative procedural law requires this as a precondition, may lodge a complaint with the supervisory authority against an implementing agency's use of a system that, in violation of these rules, has not been subject to a proper impact assessment or a proper repeat impact assessment.

2. If the supervisory authority concludes that the complaint is well founded, it may use the powers mentioned in Article 15(5)(c) and (d). The supervisory authority shall serve the complainant and the implementing authority with a reasoned decision within three months after receipt of the complaint, save for exceptional circumstances. It shall keep the complainant informed of the implementing authority's response.

3. Without prejudice to more favourable conditions, it shall be ensured that, in accordance with the applicable law, members of the public concerned having a sufficient interest, or alternatively, maintaining the impairment of a right, where

administrative procedural law requires this as a precondition:

a. have access to a review procedure before a court of law following a complaint lodged under paragraph 1. This review procedure could challenge:

i. the rejection and dismissal of a complaint by the supervisory authority;

ii. the lack of a decision of the supervisory authority within three months, save for exceptional circumstances;

iii. the continuing use of a system by an implementing authority after a recommendation of the supervisory authority to stop using the system.

b. have access to a review procedure before a court of law to challenge the legality of decisions that are reviewable in accordance with the applicable law, because they were made or supported by a system that was not subject to a proper impact assessment, or a proper repeat impact assessment.

4. What constitutes a sufficient interest and impairment of a right shall be determined consistently with the objective of giving the public concerned wide access to justice. To that end, any non-governmental organisation, meeting requirements defined by law, shall also be deemed to fulfil the requirements of paragraphs 1 or 3.

5. Any such procedure shall be fair, equitable, timely and not prohibitively expensive.

Annexes

Annex 1: Systems Always Subject to an Impact Assessment

Comment:

Annex 1 has not been drafted because the content should be left to political deliberation by the competent legislative bodies. A good starting point for the content of this Annex would be Annex III of the Draft AI Regulation (COM (2021) 206 final). Typical candidates would thus be any facial recognition systems, systems determining eligibility for social security, or certain systems for predictive policing. However, Annex III of the Draft AI Regulation still has some gaps, most obviously with regard to critical infrastructures, where rail and air traffic, as well as telecommunication networks and other digital infrastructure, are missing. In addition, systems used, for example, by tax authorities, or authorities in the fields of environmental or economic law should also be included. Other areas of public administration might raise similar concerns and therefore qualify as high risk areas.

Consequently, the Project Team encourages all respective legislative bodies to go beyond the definition of high risk AI in the Draft AI Regulation when compiling Annex 1. The criteria mentioned in Article 6 and Annexes 3 and 4 of these Model Rules can be helpful in determining what systems should be included in Annex 1. The content of the Annex should also be inspired by public debate.

A legislative option might be to provide for regular evaluation of Annexes 1 and 2. If additional ADMSs are covered by Annex 1, the legislator should also decide whether already implemented systems require an impact assessment.

Annex 1 should have two sections:

- systems that are always subject to an impact assessment – but the implementing authority must still determine whether the system is high risk;
- systems that are always high risk.

Annex 2: Systems Not Requiring an Impact Assessment

Comment:

Annex 2 has not been drafted because the content should be left to political deliberation.

Annex 2 would contain different types of systems that are already so widely established that their risks are well-known and easily manageable, and systems that are unsuitable for the high degree of public scrutiny that an impact assessment provides (mainly systems used in the area of national security).

Annex 3: Screening Questionnaire

Comment:

The questions in the screening questionnaire should be modelled after the questions in Annex 4. However, there is a crucial difference: While Annex 4 requires the implementing authority to formulate an answer, Annex 3 will mostly be a multiple-choice test modelled on the Canadian Algorithmic Impact Assessment. The results of the multiple-choice test will determine whether the system poses a low risk (ie no impact assessment necessary), a substantial risk (impact assessment necessary) or a high risk (impact assessment with public/expert participation necessary). Some questions will ask for short explanations in order to increase transparency.

The Project Team has opted not to design a complete screening questionnaire. It lacks the technical expertise and political legitimacy to assign the appropriate weight to the individual questions, which is to some extent also a political exercise.

What risk value is assigned to each answer will depend on the importance of the issue. For instance, questions pertaining to concerns about discrimination should be assigned rather high risk values. Questions about certain transparency measures that mainly help experts understand the system, but are not very relevant for the public, should have a lower risk value.

The following questions are examples as to what Annex 3 should look like:

Question type 1: Yes-no answers on the risk level

Will the algorithm used be a (trade) secret?

Yes [risk value of 1 or higher]
No [risk value of 0]

Question type 2: Several answers on a sliding scale

Is the impact resulting from the decision reversible?

☐ *Yes, it is easily reversible. [risk value of 0]*
☐ *It is likely to be reversible. [risk value of 1]*
☐ *It is difficult to reverse. [risk value of 2]*
☐ *It is irreversible. [risk value of 3 or higher]*

Individual answers might also have no exact definition, in order to emphasise that there are no clear-cut distinctions but the answers are rather a question of degree:

On a scale from 1 to 5, how easily can the impact of the decision be reversed?

(0 = very easily, 5 = not at all)

Question type 3: Risk reduction measures

Some answers – to questions of either type – could also be assigned a negative value when they indicate that the implementing authority has taken risk management measures that go beyond what can be reasonably expected.

Did you consult with the impacted communities about your definition of fairness?

☐ *Yes [risk value of -1]*
☐ *No [risk value of 0]*

Question type 4: Reference to existing impact assessments

The risk posed by a system decreases if it is already widely used. Therefore, the risk score should be significantly reduced if other authorities have already conducted an impact assessment. The reduction depends on the maximum risk score. It should have a greater weight than individual questions.

Have other authorities conducted an impact assessment under these Model Rules for the same or a similar system in a comparable context?

☐ *Yes, with the same system. [risk score reduction by 10]*
☐ *Yes, with a very similar system. [risk score reduction by 8]*
☐ *Yes, with a similar system. [risk score reduction by 4]*
☐ *Yes, in the same context. [risk score reduction by 10]*
☐ *Yes, in very similar context. [risk score reduction by 8]*
☐ *Yes, in a similar context. [risk score reduction by 4]*
☐ *No, [no risk score reduction]*

Annex 4A: Questionnaire for the Impact Assessment Report (Standard Version)

Comment:

The following questions and many formulations are based on the High-Level Expert Group's Assessment List for Trustworthy AI (ALTAI).[1] Several questions were added or modified by the ELI Project Team. Some of these changes are inspired by the European Commission's Draft AI Regulation,[2] the position paper of the German 'KI Bundesverband' on EU Regulation of AI,[3] the Council of Europe's Ad Hoc Committee on AI (CAHAI)'s Feasibility Study on AI,[4] and by the European Fundamental Rights Agency's Report 'Getting the Future Right'.[5]

The following footnotes make visible the sources for the questionnaire. A reasoned evaluation of the numerous questionnaires was outside the scope of the project.

General Remarks

Please give meaningful explanations to your answers and avoid one-word answers. If any question does not seem appropriate for your system, explain the reasons for this.

In all of the following sections, consider:

- Training measures to make those members of staff using the system aware about the relevant technical, ethical and legal issues;
- Ongoing monitoring during the use of the system;
- Suitable information for/communication with persons concerned, in particular where their cooperation is required;
- Mechanisms for persons concerned to flag any issues they encounter;
- Mechanisms to address problems that might arise during use of the system.

The structure of the questionnaire builds on Article 6.

a) a description of the purposes and operations of the system, in particular:

i. the development of the system, in particular its algorithms;

1. Who designed the system and how was it developed?[6] Who was it purchased from?

2. Are there any harmonised standards, as published in the Official Journal of the EU, or technical specifications, that (partly) apply to your system?

ii. the nature and technical characteristics of the system;

3. Describe in general, if possible in a non-technical manner, the technology/technologies you intend to use.

4. Where applicable, give a description of pre-determined or envisaged changes to the system and its performance.

iii. the selection of training, validation and testing data;

5. Describe the training methodologies and techniques used. This should include information about the provenance of those datasets, their scope and main characteristics; how the data was obtained and selected; labelling procedures and data cleaning methodologies.[7]

iv. the context in which the system is used, in particular the public objectives as defined in the applicable law;

6. What administrative task(s) does the system perform? What is its purpose? Who is responsible for its implementation, its supervision and the handling of complaints?

v. the system's interrelation with other digital systems deployed by the implementing authority or other public authorities;

---

[1] Independent High-Level Expert Group on Artificial Intelligence, Assessment List for Trustworthy AI, <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment> accessed 14 December 2021.
[2] Commission, Draft AI Regulation (COM (2021) 206 final).
[3] KI Bundesverband, Position Paper on EU-Regulation of Artificial Intelligence, <https://ki-verband.de/wp-content/uploads/2021/02/Final_Regulierung-komprimiert-1.pdf> accessed 14 December 2021.
[4] Council of Europe, Ad Hoc Committee on Artificial Intelligence, 'Feasibility Study' (CAHAI (2020) 23).
[5] EU Fundamental Rights Agency, Getting the future right, <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights> accessed 14 December 2021.
[6] Cf Commission (n 2), Annex 4 Draft AI Regulation, 2.
[7] Commission (n 2), Annex 4 Draft AI Regulation, 2 (d).

7. Does the system interact with any other hardware or software systems? If so, how?

b) an assessment of the performance, effectiveness and efficiency of the system with regard to the public objectives as defined in the applicable law, in particular whether the performance of the system will be flawed by low quality data during its use;

8. In what way will the system perform the relevant administrative task more effectively compared to the state of play (eg quicker processing, higher accuracy, lower costs)?

9. How will the system and its decisions be accepted by the public and persons concerned?

10. Have you considered the risk of the system being 'gamed' or otherwise inappropriately used? Have you put in place any measures to mitigate or avoid it?

c) an assessment of the specific and systemic impact of the system;

In considering the following questions, please take into account direct and indirect impacts, their severity, duration and reversibility.[8] Consider the impact of the system on:

i. fundamental or other individual rights or interests, in particular the rights to privacy and data protection, the right to non-discrimination and the right to good administration;

11. Is your system being trained, was it developed, or is it operated by using or processing personal data (including special categories of personal data)? If so, what are your safeguards for compliance with data protection obligations?[9]

12. Does the system potentially negatively discriminate against people on the basis of any of the following grounds (non-exhaustively): sex, gender, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, place of birth, disability, age or sexual orientation?[10]

13. If the system has an interface with the public, did you assess whether that interface reflects the variety of digital literacy and its usability by those with special needs or disabilities or those at risk of exclusion?[11]

14. Does the system protect the right to be heard, the duty to give reasons, access to the file, and other elements of the right to good administration, in accordance with the applicable law?

15. Is there any other fundamental right specifically important with regard to your particular system?

ii. democracy, societal and environmental well-being;

16. Could the system have a negative impact on election processes, public discourse, and other similarly important aspects of democracy, and did you minimise any such impact?[12]

17. Does the system control or influence critical public infrastructure (eg transport, communication or energy)?[13]

18. Is there any other impact on societal and environmental well-being (eg education, digital literacy, regional disparities, energy consumption or greenhouse gas emissions) specifically important with regard to your particular system?

iii. the administrative authority itself, in particular the estimated acceptance of the system and its decisions by the staff, the risks of over- or under-reliance on the system by the staff, and the level of digital literacy within the authority;

19. Does the system impact working conditions within the implementing authority?[14]

20. Did you ensure that staff understands how the system operates, its capabilities and limitations, in order to avoid over- or under-reliance?

[8] Criteria partly drawn from the Canadian Algorithmic Impact Assessment.
[9] Independent High-Level Expert Group on AI (n 1), 6; Council of Europe, Ad Hoc Committee on AI (n 4), 35–37.
[10] Independent High-Level Expert Group on AI (n 1), 5; Council of Europe, Ad Hoc Committe on AI (n 4), 31–33.
[11] Independent High-Level Expert Group on AI (n 1), 17.
[12] Independent High-Level Expert Group on AI (n 1), 20; Council of Europe, Ad Hoc Committee on AI (n 4), 39–41.
[13] Commission (n 2), Annex III Draft AI Regulation.
[14] Independent High-Level Expert Group on AI (n 1), 20.

d) an assessment of the measures taken to ensure:

i. maximisation of benefits to be achieved by deploying the system with regard to public objectives as defined in the applicable law;

21. Have you considered how to maximise the benefits to the public by deploying the system?

ii. minimisation of identified risks and mitigation of possible negative outcomes;

22. Have you put in place risk detection and response mechanisms, considering inter alia the minimisation of potential systemic risks? Have you established a quality management system and/or a risk management system?[15]

23. If not already mentioned above or addressed in the preceding questions: have you implemented measures in order to minimise any of the risks identified or to mitigate any negative outcomes possible?

iii. human agency, oversight and control of the system;

24. Do you adequately inform persons concerned that they are interacting with an algorithmic decision-making system?[16]

25. Do persons concerned have an alternative option to using, or being made subject to a decision by, the system?

26. What measures did you take to ensure that the system can be effectively controlled or overseen by humans? Can staff members use other means than the system to arrive at their decision?

iv. high data quality;

27. Did you put in place measures to ensure that the data used in the system is up-to-date, of high quality, complete and representative of the environment the system will be deployed in?[17]

v. accuracy across groups, precision and sensitivity;

28. Describe the measures to ensure a level of accuracy,[18] precision[19] and sensitivity[20] of the system required to avoid negative consequences.[21]

vi. technical robustness and safety; resilience to attacks; data security; fall-back plans; reliability; and reproducibility of decisions;

29. Are there sufficient safeguards against cyber-attacks, misuse, manipulation of data, malicious or inappropriate use, technical faults, defects, outages, attacks, or environmental threats?[22]

30. Did you define tested fall-back plans to address system errors, faults or inconsistencies of whatever origin (external or internal), and put governance procedures in place to trigger them?[23]

31. Did you put in place measures to evaluate and ensure the system's reliability and reproducibility?[24]

vii. transparency of the system and explainability of its decisions;

32. How will you inform persons concerned and the public about the existence and functioning of the system?[25]

33. Can you explain the decision(s) of the system to the persons concerned?[26]

---

[15] Commission (n 2), Art. 9 and 17 Draft AI Regulation.
[16] Independent High-Level Expert Group on AI (n 1), 7.
[17] Independent High-Level Expert Group on AI (n 1), 10.
[18] Accuracy means the number of correctly predicted data points out of all the data points, ie the number of true positives and true negatives divided by the number of true positives, true negatives, false positives, and false negatives; <https://deepai.org/machine-learning-glossary-and-terms/accuracy-error-rate> accessed 14 December 2021.
[19] The fraction of relevant instances among all retrieved instances, ie the number of true positives divided by the sum of true and false positives, <https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall> accessed 14 December 2021.
[20] Defined as the fraction of retrieved instances among all relevant instances, ie the number of true positives divided by the sum of true positives and false negatives; <https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall> accessed 14 December 2021.
[21] Independent High-Level Expert Group on AI (n 1), 10.
[22] Independent High-Level Expert Group on AI (n 1), 10.
[23] Independent High-Level Expert Group on AI (n 1), 11.
[24] Independent High-Level Expert Group on AI (n 1), 11.
[25] Independent High-Level Expert Group on AI (n 1), 15.
[26] Independent High-Level Expert Group on AI (n 1), 15.

viii. traceability in order to enable the monitoring of the system's operations;

34. Did you put in place measures that address the traceability of the system during its entire lifecycle (eg logging of the system's processes and outputs)?[27]

ix. accountability, in particular oversight, auditability, clear allocation of responsibilities, self-monitoring, benchmarking, and the possibility of redress for injury or harm caused by the system;

35. Did you establish mechanisms that facilitate the system's auditability (eg documentation of the development process, the sourcing of training data and complaints about negative impacts, and the logging of the system's processes)?[28]

36. Have you assigned clear responsibilities for every stage of the system (eg development, deployment, use, oversights, handling of complaints, and fixing of errors)?

e) unless the system is listed as 'always high risk' in Annex 1, a concluding determination of the risk level;

f) an overall assessment of the necessity and proportionality of the processing operations in relation to the purposes, in particular trade-offs between different factors set out in this Article and reasonable alternatives to the project;

g) a reasoned statement on the legality of the use of the system under the applicable law, in particular data protection law, administrative procedure law and applicable sectoral legislation.

[27] Independent High-Level Expert Group on AI (n 1), 14.
[28] Independent High-Level Expert Group on AI (n 1), 21.

Annex 4B: Questionnaire for the Impact Assessment Report (Extended Version)

Comment:

The following questions and many formulations are based on the High-Level Expert Group's Assessment List for Trustworthy AI (ALTAI).[29] Several questions were added or modified by the ELI Project Team. Some of these changes are inspired by the European Commission's Draft AI Regulation,[30] the position paper of the German 'KI Bundesverband' on EU Regulation of AI,[31] the Council of Europe's Ad Hoc Committee on AI (CAHAI)'s Feasibility Study on AI,[32] and by the European Fundamental Rights Agency's Report 'Getting the Future Right'.[33]

General Remarks

Please give meaningful explanations to your answers and avoid one-word answers. If any question does not seem appropriate for your system, explain the reasons for this.

In all of the following sections, consider

- Training measures to make those members of staff using the system aware about the relevant technical, ethical and legal issues;
- Ongoing monitoring during the use of the system;
- Suitable information for/communication with persons concerned, in particular where their cooperation is required;
- Mechanisms for persons concerned to flag any issues they encounter;
- Mechanisms to address problems that might arise during the use of the system.

The structure of the questionnaire builds on Article 6.

a) a description of the purposes and operations of the system, in particular:

i. the development of the system, in particular its algorithms;

1. Who designed the system and how was it developed?[34] Who was it purchased from?

   1.1 How did you select the developer/manufacturer? What was their edge over competitors?

   1.2 Has the system been put on the market/into service elsewhere?

   1.3 What methods and steps were used in the development of the system? Were pre-trained systems or tools provided by third parties? If so, did you modify them?[35]

   1.4 What computational resources were used to develop, train, test and validate the system?

2. Are there any harmonised standards, as published in the Official Journal of the EU, or technical specifications, that (partly) apply to your system?

ii. the nature and technical characteristics of the system;

3. Describe in general, if possible in a non-technical manner, the technology/technologies you intend to use.

   3.1. This should include[36]: (1) The date and the version of the system; (2) A description of hardware on which the system is intended to run; (3) Where the system is a component of products, photographs or illustrations showing external features, marking and internal layout of these products.

   3.2 Technically speaking, what are the characteristics of the system? In what ways does it work? Describe the system architecture and explain how software components build on or feed into each other and integrate into the overall processing.

   For example,[37]

   3.2.1 Is it rules-based, applying clear 'if-then rules'?

[29] Independent High-Level Expert Group on Artificial Intelligence (n 1).
[30] Commission, Draft AI Regulation (COM (2021) 206 final).
[31] KI Bundesverband (n 3).
[32] Council of Europe, Ad Hoc Committee on Artificial Intelligence (n 4).
[33] EU Fundamental Rights Agency (n 5).
[34] Cf Commission (n 2), Annex 4 Draft Regulation, 2.
[35] Cf Commission (n 2), Annex 4 Draft AI Regulation, 2 (a).
[36] Cf Annex 4 Commission (n 2), Regulation, 1 (c–f).
[37] EU Fundamental Rights Agency (n 5), 27; KI Bundesverband, Position Paper on EU-Regulation of Artificial Intelligence, 8.

3.2.2 Does it rely on more traditional statistical methods to find correlations, eg regression analysis?

3.2.3 Does it use logic- or knowledge-based approaches?

3.2.4 Is it a self-learning/machine learning algorithm, using eg supervised, unsupervised or reinforcement learning?

3.2.5 Is the model trained only once or is it continuously retrained?

3.2.6 Does it apply deep learning?

3.3 What is the degree of automation of the system?[38] Does the system make final decisions or does it only make recommendations to humans?

3.4 What are the design specifications? What is the general logic of the system and its algorithms?

3.5 What were the key design choices, including the rationale and assumptions made (eg about the persons on which the system is intended to be used)?

3.6 What were the main classification choices?

3.7 What is the system designed to optimise for?

3.8 What is the relevance of the different parameters?

3.9 What were the key trade-offs regarding the technical solutions adopted?

4. Where applicable, give a description of pre-determined or envisaged changes to the system and its performance.

iii. the selection of training, validation and testing data;

5. Describe the training methodologies and techniques used. This should include information about the provenance of those datasets, their scope and main characteristics; how the data was obtained and selected; labelling procedures and data cleaning methodologies.[39]

In particular, describe what training data you have used to train the system.

5.1 Is the dataset static (ie fixed and clearly defined) or is it dynamic (ie continuously fed with new data)?

5.2 Is the data and the data generating process open or under control?

5.3 Is the system trained on personal data or on neutral data?

iv. the context in which the system is used, in particular the public objectives as defined in the applicable law;

6. What administrative task(s) does the system perform? What is its purpose? Who is responsible for its implementation, its supervision and the handling of complaints?

6.1 What is your motivation for using this system?

6.2 Does your public authority have prior experience with similar systems (ie similar technologies)?

v. the system's interrelation with other digital systems deployed by the implementing authority or other public authorities;

7. Does the system interact with any other hardware or software systems? If so, how?

b) an assessment of the performance, effectiveness and efficiency of the system with regard to the public objectives as defined in the applicable law, in particular whether the performance of the system will be flawed by low quality data during its use;

8. In what way will the system perform the relevant administrative task more effectively compared to the state of play (eg quicker processing, higher accuracy, lower costs)?

9. How will the system and its decisions be accepted by the public and persons concerned?

9.1 Does your system rely on the initiative, cooperation or trust of persons concerned? If so, how do you ensure their cooperation? Have you put in place incentives for using the system?

10. Have you considered the risk of the system being 'gamed' or otherwise inappropriately used? Have

[38] EU Fundamental Rights Agency (n 5) 27.
[39] Commission (n 2), Annex 4 Draft AI Regulation, 2 (d).

you put in place any measures to mitigate or avoid it?

c) an assessment of the specific and systemic impact of the system;

In considering the following questions, please take into account direct and indirect impacts, their severity, duration and reversibility.[40] Consider the impact of the system on:

i. fundamental or other individual rights or interests, in particular the rights to privacy and data protection, the right to non-discrimination and the right to good administration;

11. Is your system being trained, was it developed, or is it operated by using or processing personal data (including special categories of personal data)? If so, what are your safeguards for compliance with data protection obligations?[41]

    11.1 Did you consider the implications of the system's non-personal training-data or other processed non-personal data for privacy or secrecy interests of legal persons?[42]

    11.2 In which ways does your system follow the concept of privacy by design and default?[43]

12. Does the system potentially negatively discriminate against people on the basis of any of the following grounds (non-exhaustively): sex, gender, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, place of birth, disability, age or sexual orientation?[44]

    12.1 Did you consider diversity and representativeness of persons concerned in the data?[45]

    12.2 Did you test for specific target groups or problematic use cases?[46]

    12.3 Do you have a definition of fairness? Is your definition of fairness commonly used and implemented in any phase of the process of setting up the system?

    12.4 Did you establish mechanisms to ensure fairness in your system?

13. If the system has an interface with the public, did you assess whether that interface reflects the variety of digital literacy and its usability by those with special needs or disabilities or those at risk of exclusion?[47]

    13.1 Did you involve or consult with persons concerned in need of assistive technology during the planning and development phase of the system?[48]

    13.2 Did you ensure that Universal Design Principles[49] and accessibility standards[50] are considered during every step of the planning and development process, if applicable?[51]

14. Does the system protect the right to be heard, the duty to give reasons, access to the file, and other elements of the right to good administration, in accordance with the applicable law?

15. Is there any other fundamental right specifically important with regard to your particular system?

ii. democracy, societal and environmental well-being;

16. Could the system have a negative impact on election processes, public discourse, and other similarly important aspects of democracy, and did you minimise any such impact?[52]

17. Does the system control or influence critical public

[40] Criteria partly drawn from the Canadian Algorithmic Impact Assessment.
[41] Independent High-Level Expert Group on AI (n 1), 6; Council of Europe, Ad Hoc Committee on AI (n 4), 35–37.
[42] Independent High-Level Expert Group on AI (n 1), 13.
[43] Independent High-Level Expert Group on AI (n 1), 12.
[44] Independent High-Level Expert Group on AI (n 1), 5; Council of Europe, Ad Hoc Committee on AI (n 4), 31–33.
[45] Independent High-Level Expert Group on AI (n 1), 16.
[46] Independent High-Level Expert Group on AI (n 1), 16.
[47] Independent High-Level Expert Group on AI (n 1), 17.
[48] Independent High-Level Expert Group on AI (n 1), 17.
[49] <www.cen.eu/news/brief-news/Pages/NEWS-2019-014.aspx> accessed 14 December 2021.
[50] <www.iso.org/standard/58625.html; www.iso.org/standard/33987.html> accessed 14 December 2021; <www.iso.org/obp/ui/#iso:std:iso:9241:-171:ed-1:v1:en> accessed 14 December 2021; <http://mandate376.standards.eu/standard> accessed 14 December 2021.
[51] Independent High-Level Expert Group on AI (n 1), 17.
[52] Independent High-Level Expert Group on AI (n 1), 20; CAHAI (2020) 23, 39–41.

infrastructure (eg transport, communication or energy)?[53]

18. Is there any other impact on societal and environmental well-being (eg education, digital literacy, regional disparities, energy consumption or greenhouse gas emissions) specifically important with regard to your particular system?

iii. the administrative authority itself, in particular the estimated acceptance of the system and its decisions by the staff, the risks of over- or under-reliance on the system by the staff, and the level of digital literacy within the authority;

19. Does the system impact working conditions within the implementing authority?[54]

    19.1 Could the system create the risk of de-skilling your staff? Did you take measures to counteract de-skilling risks?[55]

    19.2 Does the system promote or require new (digital) skills? Did you provide training opportunities and materials for re- and up-skilling?[56]

    19.3 Does the system use data with security classification? What special measures are in place to protect this data against unauthorised access by staff members?

    19.4 Does the system reduce the number of staff required in your administrative agency, or does it require special abilities not yet available in your existing staff?

    19.5 Did you pave the way for the introduction of the system in your organisation by informing and consulting with impacted staff and their representatives in advance?[57]

20. Did you ensure that staff understands how the system operates, its capabilities and limitations, in order to avoid over- or under-reliance?

d) an assessment of the measures taken to ensure:

i. maximisation of benefits to be achieved by deploying the system with regard to public objectives as defined in the applicable law;

21. Have you considered how to maximise the benefits to the public by deploying the system?

    21.1 Have you considered possibilities for innovative fulfilment of the administrative objective through exploiting the potentials of the system?

ii. minimisation of identified risks and mitigation of possible negative outcomes;

22. Have you put in place risk detection and response mechanisms, considering inter alia the minimisation of potential systemic risks? Have you established a quality management system and/or a risk management system?[58]

23. If not already mentioned above or addressed in the preceding questions: have you implemented measures in order to minimise any of the risks identified or to mitigate any negative outcomes possible?

iii. human agency, oversight and control of the system

24. Do you adequately inform persons concerned that they are interacting with an algortihmic decision-making system?[59]

    24.1 Could the system manipulate actions or reactions of persons concerned, eg by nudging?

25. Do persons concerned have an alternative option to using, or being made subject to a decision by, the system?

26. What measures did you take to ensure that the system can be effectively controlled or overseen by humans? Can staff members use other means than the system to arrive at their decision?

    26.1 Is it controlled or overseen by a Human-in-the-Loop, Human-on-the-Loop or Human-in-Command?[60]

[53] Commission (n 2), Annex III Draft AI Regulation.
[54] Independent High-Level Expert Group on AI (n 1), 20.
[55] Independent High-Level Expert Group on AI (n 1), 20.
[56] Independent High-Level Expert Group on AI (n 1), 20.
[57] Independent High-Level Expert Group on AI (n 1), 20.
[58] Commission (n 2), Articles 9 and 17 Draft AI Regulation.
[59] Independent High-Level Expert Group on AI (n 1), 7.
[60] Independent High-Level Expert Group on AI (n 1), 8.

26.2 Is the requirement of human oversight already built into the system, or do you need to take specific organisational measures?[61]

26.3 In particular, have the humans who oversee the system been given specific training on how to exercise oversight?[62]

26.4 Does this training ensure that they

a. understand the capacities and limitations of the system?

b. understand the risk of 'automation bias'?

c. can correctly interpret the system's output?

d. are given criteria or instructions when not to use or rely on the system?

26.5 Is there a 'stop button' or procedure to safely abort an operation when needed?[63]

26.6 Did you take any specific oversight and control measures to reflect the self-learning or autonomous nature of the system?[64]

iv. high data quality;

27. Did you put in place measures to ensure that the data used in the system is up-to-date, of high quality, complete and representative of the environment the system will be deployed in?[65]

27.1 Do the training, validation and testing datasets consider, to the extent required by the system's purpose, the specific geographical, behavioural or functional setting within which the system is used?[66]

27.2 Is there a risk of high data dispersion with extreme outliers that might skew the algorithm? If yes, what measures did you take against this risk?

27.3 Did you establish appropriate data governance and management practices such as design choices; data collection; data preparation processing operations (eg annotation, labelling, cleaning, enrichment and aggregation); the formulation of relevant assumptions (in particular what information the data is supposed to measure and represent); assessments of the availability, quantity and suitability of the datasets that are needed; examination of possible biases; identification of possible data gaps or shortcomings and means to address them?[67]

v. accuracy across groups, precision and sensitivity;

28. Describe the measures to ensure a level of accuracy,68 precision[69] and sensitivity[70] of the system required to avoid negative consequences. [71]

28.1 Did you consider whether the system's operation can invalidate the data or assumptions it was trained on, and how this might lead to adverse effects?[72]

28.2 Did you put processes in place to ensure that the level of accuracy and precision of the system to be expected by persons concerned is properly communicated?[73]

28.3 Did you declare the level of accuracy and the relevant metrics in the instructions of use, or make them available in another manner?[74]

[61] Commission (n 2), Article 14 (1) and (3) Draft AI Regulation.
[62] Independent High-Level Expert Group on AI (n 1), 8.
[63] Independent High-Level Expert Group on AI (n 1), 8.
[64] Independent High-Level Expert Group on AI (n 1), 8.
[65] Independent High-Level Expert Group on AI (n 1), 10.
[66] Commission (n 2), Article 10 (4) Draft AI Regulation.
[67] Compare Commission (n 2), Article 10 (2) Draft AI Regulation.
[68] Accuracy means the number of correctly predicted data points out of all the data points, ie the number of true positives and true negatives divided by the number of true positives, true negatives, false positives, and false negatives; <https://deepai.org/machine-learning-glossary-and-terms/accuracy-error-rate> accessed 14 December 2021.
[69] The fraction of relevant instances among all retrieved instances, ie the number of true positives divided by the sum of true and false positives <https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall> accessed 14 December 2021.
[70] Defined as the fraction of retrieved instances among all relevant instances, ie the number of true positives divided by the sum of true positives and false negatives; <https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall> accessed 14 December 2021.
[71] Independent High-Level Expert Group on AI (n 1), 10.
[72] Independent High-Level Expert Group on AI (n 1), 10.
[73] Independent High-Level Expert Group on AI (n 1), 10.
[74] Commission (n 2), Article 15 (2) Draft AI Regulation.

vi. technical robustness and safety; resilience to attacks; data security; fall-back plans; reliability; and reproducibility of decisions;

29. Are there sufficient safeguards against cyber-attacks, misuse, manipulation of data, malicious or inappropriate use, technical faults, defects, outages, attacks, or environmental threats?[75]

    29.1 Could the system have adverse, critical or damaging effects (eg to human or societal safety) if such risks materialise?

    29.2 Is the system compliant with general or specific cybersecurity standards? Is it certified for cybersecurity (eg the certification scheme created by the Cybersecurity Act in Europe[76])?[77]

    29.3 Did you put measures in place to ensure the integrity, robustness and overall security of the system against potential attacks over its lifecycle that are appropriate to the risks and circumstances?[78] Did you pentest these measures?

    29.4 What length is the expected timeframe within which you provide security updates for the system?[79]

30. Did you define tested fall-back plans to address system errors, faults or inconsistencies of whatever origin (external or internal), and put governance procedures in place to trigger them?[80]

31. Did you put in place measures to evaluate and ensure the system's reliability and reproducibility?[81]

    31.1 Did you test whether specific contexts or conditions need to be considered to ensure reproducibility?

    31.2 Did you assess the dependency of the system's decisions on its stable and reliable behaviour?

    31.3 Did you align the reliability/testing requirements to the appropriate levels of stability and reliability?

vii. transparency of the system and explainability of its decisions;

32. How will you inform persons concerned and the public about the existence and functioning of the system?[82]

33. Can you explain the decision(s) of the system to the persons concerned?[83]

    33.1 Do you continuously survey the persons concerned to determine whether they understand the decision(s) of the system?[84]

viii. traceability in order to enable the monitoring of the system's operations;

34. Did you put in place measures that address the traceability of the system during its entire lifecycle (eg logging of the system's processes and outputs)?[85]

    34.1 Can you trace back which algorithms, rules and/or data were used by the system to make a certain decision(s) or recommendation(s)?[86]

    34.2 Did you put in place measures to continuously assess the quality of the output(s) of the system?[87]

    34.3 Did you put adequate logging practices in place to record the decision(s) or recommendation(s) of the system? Are these logs kept for an appropriate time?[88]

[75] Independent High-Level Expert Group on AI (n 1), 10.
[76] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification.
[77] Independent High-Level Expert Group on AI (n 1), 9.
[78] Independent High-Level Expert Group on AI (n 1), 9.
[79] Independent High-Level Expert Group on AI (n 1), 9.
[80] Independent High-Level Expert Group on AI (n 1), 11.
[81] Independent High-Level Expert Group on AI (n 1), 11.
[82] Independent High-Level Expert Group on AI (n 1), 15.
[83] Independent High-Level Expert Group on AI (n 1), 15.
[84] Independent High-Level Expert Group on AI (n 1), 15.
[85] Independent High-Level Expert Group on AI (n 1), 14.
[86] Independent High-Level Expert Group on AI (n 1), 14.
[87] Independent High-Level Expert Group on AI (n 1), 14.
[88] Commission (n 2), Article 20, 29 (5) Draft AI Regulation.

ix. accountability, in particular oversight, auditability, clear allocation of responsibilities, self-monitoring, benchmarking, and the possibility of redress for injury or harm caused by the system;

35. Did you establish mechanisms that facilitate the system's auditability (eg documentation of the development process, the sourcing of training data and complaints about negative impacts, and the logging of the system's processes)?[89]

36. Have you assigned clear responsibilities for every stage of the system (eg development, deployment, use, oversights, handling of complaints, and fixing of errors)?

    36.1 Did you consider establishing an AI ethics review board or a similar mechanism to discuss the overall accountability and ethics practices, including potential unclear grey areas?[90]

    36.2 Have you considered introducing benchmarking procedures to compare the performance and risks of the system's output with human-made decisions in the same areas?

    36.3 Did you establish a process for third parties (eg suppliers, persons concerned, distributors/ vendors, workers or civil society organisations) to report potential vulnerabilities, risks or biases in the system? Does this process foster revision of the risk management process?[91]

    36.4 For systems that can adversely affect individuals, have self-monitoring mechanisms (in particular redress by design mechanisms) been put in place?[92]

e) unless the system is listed as 'always high risk' in Annex 1, a concluding determination of the risk level;

    37.1 If the system was not listed in Annex 1: After completing the questions of this Annex in detail, would you change your risk evaluation according to Annex 3? In particular, would you now think that your system poses a high risk?

    37.2 If the system was listed in Annex 1 as 'substantial risk': Please complete the questionnaire in Annex 3 using the information gained during the impact assessment. According to the questionnaire, does it constitute a high risk?

f) an overall assessment of the necessity and proportionality of the processing operations in relation to the purposes, in particular trade-offs between different factors set out in this Article and reasonable alternatives to the project;

    38.1 What would the alternatives to using this system be?

    a. If possible, did you consider the use of different systems or solutions not using algorithmic decision-making?

    b. Why did you ultimately decide to use this system?

    38.2 In your view, why is it acceptable to take the risks associated with the use of the system?

g) a reasoned statement on the legality of the use of the system under the applicable law, in particular data protection law, administrative procedure law and applicable sectoral legislation.

[89] Independent High-Level Expert Group on AI (n 1), 21.
[90] Independent High-Level Expert Group on AI (n 1), 22.
[91] Independent High-Level Expert Group on AI (n 1), 22.
[92] Independent High-Level Expert Group on AI (n 1), 22.

ELI Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration with Comments and Sources

Article 1: Purpose and Scope

Sources

Article 1(1) and (4) Environmental Impact Assessment Directive 2011/92/EU

Comments

(1.1.) Paragraph 1 defines the purpose of the Model Rules. It is important to stress that the Model Rules provide for an impact assessment procedure, not for (new) material standards. The relevant 'impacts on the public' as mentioned in paragraph 1 can be taken from the list in Article 6(2)(c) and Annex 4.

(1.2.) Paragraph 2 defines when an impact assessment is necessary. Annex 1 is to contain a list of systems for which an impact assessment is always necessary. Which systems are listed in this Annex is mainly a political decision. For this reason, the Project Team refrained from drafting this Annex. States that wish to implement the Model Rules might consider including systems that are subject to intense political debate, or that are generally considered very risky. A good starting point for the content of this Annex would be Annex III of the Draft AI Regulation (COM (2021) 206 final). The Project Team decided against a direct reference to this Regulation for two reasons: First, non-EU countries might also wish to implement the Model Rules. Second, Annex III of the Draft Regulation appears too narrow in some regards (eg its definition of critical infrastructure is not in line with already existing EU law on this matter (NIS Directive 2016/1148) and does not include for instance telecommunication or rail traffic). The Project Team encourages Member States to go beyond the definition of high risk AI in the Draft AI Regulation when compiling Annex 1. The Draft AI Regulation applies to both the public and the private use of AI and must therefore avoid excessive burdens on private businesses. In contrast, these Model Rules only apply to ADMSs in the public sector, which should be subject to a higher degree of scrutiny than AI systems in general.

(1.3.) Systems not covered by Annex 1 or paragraph 3 must be subject to the screening procedure (Article 4) to determine whether an impact assessment will be necessary.

(1.4.) Systems in Annex 2 are not subject to an impact assessment. The Project Team leaves the content of Annex 2 to political decision-making. Annex 2 would contain different types of systems: systems that are obviously low risk (such as common chatbots), systems that are already so widely established that their risks are well-known and easily manageable, and systems that are unsuitable for the high degree of public scrutiny that an impact assessment provides (mainly systems used in the area of national security). The latter option should be used sparingly because Article 8 of the Model Rules allows for the protection of secrets. Annex 2 should also include systems that are used as components of products that are already subject to extensive safety regulation if such safety regulation addresses the specific risks of algorithmic decision-making. Annex II of the Draft AI Regulation provides a number of examples for such systems, especially driverless cars. The Project Team assumes that the risks of such systems are already addressed comprehensively by the specific legislation. An exception might be needed if such products will be used by public authorities in a manner which differs significantly from common practices as presumed by general product safety rules.

(1.5.) The mode of revision of annexes is to be adapted to each legal system. For instance, in the case of a directive, it may be necessary to give national parliaments the right to amend the annexes.

(1.6.) Paragraph 4 provides for a narrow emergency exception that permits the postponement of the assessment until after the deployment of the system. The clause covers emergencies such as the Covid-19 pandemic. In such cases, the need for a speedy reaction (eg for contact tracing or vaccination appointments) outweighs concerns about transparency and detailed

risk management. The paragraph does not apply to general threats of crime or terrorism.

(1.7.) The implementing authority must carry out the impact assessment before the deployment of the system. It is not possible to mandate a specific time for the impact assessment. It should accompany the development of the system. The impact assessment should start early enough to allow for meaningful changes if the implementing authority discovers problems during the assessment. However, starting too early often means that the specifics of the system are not yet known. The exact timing of the impact assessment also depends on whether the implementing authority develops the system itself or purchases it from an external provider. In the latter case, cooperation with the provider is particularly important for a successful impact assessment. Rules on this cooperation are set out in Article 7.

Article 2: Definitions

Sources

Article 1(2) Environmental Impact Assessment Directive 2011/92/EU; Article 3 (9) INSPIRE Directive 2007/2/EC; Appendix A, Directive on Automated Decision-Making (Canada); Article 2 (4) Washington Senate Bill 5116; Article III-2 (1) ReNEUAL Model Rules

Comments

Algorithmic Decision-Making System

(2.1.) The Project Team preferred the term 'algorithmic decision-making system' (ADMS) to the notion of 'artificial intelligence', because it is broader and technologically neutral. ADMS are not limited to AI or machine learning systems. Whether or not an ADMS uses AI technology is often disputable and more conventional algorithmic systems can also pose relevant risks. The Project Team is aware that the European Commission chose the term 'artificial intelligence' for its Draft AI Regulation. The Project Team considers the term 'algorithmic decision-making' more appropriate at least for public administration. Public administration does not only use highly complex systems that would commonly be described as AI, but also simpler systems where disagreement could arise whether these systems are really AI. In order to avoid such controversy, the Project Team opted for the more neutral term 'algorithmic decision-making'. Unlike in data protection law, it is irrelevant whether the system processes personal data or not.

(2.2.) The term ADMS is a broad one and encompasses a large number of computer systems, even those that are not particularly risky. However, far fewer systems will be subject to an impact assessment, because of the screening procedure prescribed by Article 4. Some ADMS that are very common and obviously not risky (such as antivirus software, automatic spell-checkers, etc) should be included in Annex 2, because they do not even warrant a screening procedure. Public authorities are always free to conduct impact assessments for systems that are not covered by the Model Rules.

(2.3.) The definition includes systems that support human decision-making. Even if there is a human making the final decision, systems that prepare or analyse data for the decision or even make specific proposals can exercise a significant influence on the decision-maker and should therefore not go unchecked.

Public Authority

(2.4.) The definition is taken from Article 3(9) of the INSPIRE Directive 2007/2/EC. It covers all levels of public administration. By its letter 'c', the definition also includes private actors exercising administrative functions.

(2.5.) The judiciary is not covered by the definition, unless it acts in an administrative capacity. The Project Team supports impact assessments for ADMS in the judiciary. The impact of such systems (eg systems that predict recidivism in criminals) can be at least as high as the impact of systems used by the executive. However, the independence of the judiciary would conflict with parts of the Model Rules (especially the competences of the supervisory authority under Article 15). Legislators wishing to introduce impact assessment for the judiciary could use these Model Rules as a blueprint, but should adapt them carefully.

Decision

(2.6.) The term 'decision' is broader than typical definitions of administrative decision in national administrative procedure or administrative justice legislation (compare, eg, Article III-2 (1) of the ReNEUAL Model Rules). An action in the sense of Article 2 does not need to be legally binding. The term also covers purely factual actions (eg, warnings or advice). A decision in the sense of these Model

Rules is also not restricted to an individual case. The definition also encompasses a change of policy (eg, to establish surveillance cameras in a certain area where an algorithmic system has detected an increase in crime). It further encompasses declaratory decisions stating that a certain fact is at hand, such as grading decisions finding someone eligible for health insurance. The definition includes actions of administrative authorities under both public and private law. It is not necessary that the affected individual realises that an action has been taken, so eg secret surveillance is covered by the definition. The term 'determination' is not limited to the binary question whether or not to act. It also includes the question how to act (eg whether to impose a strict or a lenient sanction).

Public

(2.7.) The definition of the public is relevant for public participation according to Article 11. It is drawn from Article 1(2)d of the Environmental Impact Assessment Directive (Directive 2011/92/EU), which establishes a comparable public participation mechanism.

System Provider/Data Provider

(2.8.) This definition is relevant for Article 7, which regulates the cooperation of the implementing authority and these actors. It is different from the definition of a provider in Article 3(2) of the Draft AI Regulation (COM (2021) 206 final), which focuses on the entity that places a system on the EU market, or puts it into service. In contrast, a system (or data) provider can be anyone who provides parts of the system or relevant data, regardless of their position in the supply chain. It could also be a distributor in the sense of Article 3(7) of the Draft AI Regulation. The definition of data provider also differs slightly from the definition of 'data supplier' in Principle 3(1) (n) of the ELI Principles for a Data Economy. That definition refers to any party who supplies data to another party, or undertakes to do so. Within these Model Rules, data must be supplied specifically to the implementing authority, and merely undertaking that supply is not enough.

Implementing Authority

(2.9.) The implementing authority is the main addressee of the obligations created by these Model Rules. The definition focuses on the authority that uses or intends to use the system. However, if a superior public authority (eg a ministry of finance) decides that subordinate authorities (eg local tax offices) have to use the system, the superior authority becomes responsible for the impact assessment. It can then decide to conduct the impact assessment itself – if needed, with support by one of the subordinate authorities – or even to delegate it to one of them.

Article 3: Coordination with Other Procedures

Sources

Articles 9, 35 Draft AI Regulation (Commission proposal; COM (2021) 206 final); Article 35 General Data Protection Regulation (EU) 2016/679

Comments

(3.) It is important to ensure that there is coordination between the procedures of these Model Rules and procedures provided by other legislation in order to avoid duplicating assessments. For example, it is clear that the impact assessment provided for in Article 35 of the GDPR (Regulation (EU) 2016/679) could at least be part of the information required for the Article 6 report. Similarly, EU product safety legislation imposes conformity assessments that could already address some of the issues to be developed in the Article 6 report. In particular, the current Draft Regulation of the EU on Artificial Intelligence (COM (2021) 206 final) may also, if adopted, impose some evaluations which could be of direct use for completing the Article 6 report. This is particularly true for the conformity assessment (Article 43) and the risk management system (Article 9). A priori, each legal system should, by implementing these Model Rules, identify existing procedures that can partially meet the requirements of the Article 6 report.

It is not possible to describe all the situations in which coordination between procedures is necessary, especially in the case of conflicting provisions, but the attention of the legislator is drawn to this point.

Article 4: Screening

Sources

Article 4 and Annex III Environmental Impact Assessment Directive 2011/92/EU; Article 35(1) General Data Protection Regulation (EU) 2016/679; Number 6.1.1 Directive on Automated Decision-Making (Canada)

Comments

(4.1.) Article 4 provides for a screening procedure to determine whether systems not listed in Annexes 1 or 2 will be subject to an impact assessment. The concept of screening or initial evaluation is known from the environmental and data protection impact assessment. The screening procedure does not need to be as thorough and detailed as a proper impact assessment.

(4.2.) A detailed questionnaire for the screening procedure is provided in Annex 3. Its general idea is based on the Canadian Algorithmic Impact Assessment. Most of the questions are multiple-choice questions. Each answer should be assigned a certain risk value. The sum of all answers yields a risk score which determines the risk level. Some answers that indicate that risk management measures have been implemented should be assigned negative values. The Project Team did not assign definite values to each answer as it lacks practical experience with public sector AI systems. Governments wishing to implement the Model Rules should consult experts and practitioners to assign definite values.

(4.3.) The screening procedure can have three different outcomes, described by three different levels of risk:

- Low risk, where an impact assessment is not necessary;
- Substantial risk, where an impact assessment is necessary, but Chapter 3 does not apply;
- High risk, where an impact assessment including public and expert consultations under Chapter 3 is necessary.

(4.4.) 'High risk' in these Model Rules is not the same as 'high risk' in the EU's Draft AI Regulation, although there may be overlaps. The Draft AI Regulation assigns this label based on products in which the AI system is integrated, or areas in which AI systems are used. In contrast, these Model Rules rely on a case-by-case assessment considering the specific context in which the system is used.

(4.5.) If an impact assessment is necessary, the implementing authority will have to review the initial risk evaluation after it has completed the impact assessment report (Article 6 paragraph 2 lit f). This ensures that the risk evaluation considers the results of a detailed investigation.

(4.6.) According to paragraph 2, the implementing authority shall publish the results of the screening procedure. This provides a certain level of transparency even for systems for which no impact assessment proper is necessary. As the initial questionnaire is a multiple-choice test, the implementing authority does not need to explain the results of the screening procedure. However, adding an explanation can be useful to provide additional transparency, for instance where the system only narrowly misses a risk threshold.

(4.7.) As it cannot be excluded that the screening results contain secrets, paragraph 2 concludes with a reference to Article 13(3) which is the central provision for this issue.

Article 5: Scoping

Sources

Article 5(2) Environmental Impact Assessment Directive 2011/92/EU; Article 4(2) Regulation (EU) 182/2011; § 15 *Umweltverträglichkeitsprüfungsgesetz* (German Environmental Impact Assessment Act)

Comments

(5.1.) A scoping is not mandatory, but helps the implementing authority to determine the most important issues that will have to be covered in the Article 6 report. As individual implementing authorities might lack expertise on AI, it is especially important to consult experts or, if relevant, the supervisory authority under Article 16.

(5.2.) Paragraph 2 allows the implementing authority to ask the supervisory authority to carry out the scoping. This enables the implementing authority to make the best use of the supervisory authority's expertise. Paragraph 2 also clarifies that the implementing authority must take utmost account of the supervisory authority's scoping results. The formulation 'taking utmost account of' is typical in EU administrative law, eg Article 4 (2) of Regulation (EU) 182/2011 governing comitology.

(5.3.) An important part of the scoping exercise is to clarify on which issues the assessment should focus. Typically, not all of the many aspects listed in Article 6 will require thorough examination. However, the scoping is only about setting priorities, not about excluding from the impact assessment aspects mentioned in Article 6.

Article 6: Impact Assessment Report

Sources

Articles 6-20 EU Draft AI Regulation (Commission proposal; COM (2021) 206 final); Article 35(1) and (7) General Data Protection Regulation (EU) 2016/679; Article 5(3) Environmental Impact Assessment Directive 2011/92/EU; EU High-Level Expert Group on Artificial Intelligence's Assessment List for Trustworthy AI (2020); Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the Human Rights Impacts of Algorithmic Systems (Council of Europe); 'The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law' (Report by Catelijne Muller; CAHAI (2020) 06; Feasibility Study (Council of Europe, Ad Hoc Committee on Artificial Intelligence, CAHAI (2020) 23; 'Human Rights, Democracy and Rule of Law Impact Assessment of AI Systems' (Council of Europe, Ad Hoc Committee on Artifical Intelligence, CAHAI-PDG (2021) 05; Impact Assessment provided by the Directive on Automated Decision-Making (Canada); § 16 *Umweltverträglichkeitsprüfungsgesetz* (German Environmental Impact Assessment Act)

Comments

(6.1.) Article 6 draws from various documents on legal and ethical requirements to set up an extensive and detailed list of issues to be assessed. The Project Team aimed at an Article that is more specific than most documents on AI ethics, but broad enough to cover all the legal and ethical issues that are typically discussed concerning ADMSs. In addition, the Project Team aimed at a clearly justiciable 'legal' structure for the various assessment criteria. The broad scope of the Article is justified by the large variety of impacts that ADMSs can have and of which administrative staff is not always aware. Thus, it is important to provide relatively detailed guidance on the relevant aspects. In contrast to most impact assessment proposals and Article 35 GDPR, the draft does not focus on risks alone, but also asks how the implementing authority can optimise the benefits of the systems. The assessment criteria are not binding standards that any system must fulfil: the Model Rules provide for an impact assessment, not for material standards. If an implementing authority cannot or does not want to take steps to meet a criterion in this Article, it can explain its reasons in the report. For instance, a configuration that ensures maximum transparency of a system might not be technically feasibly or too expensive. It is then up to supervisory authorities, the courts and/or the public to decide whether the use of the system is (legally or politically) acceptable.

(6.2.) While Article 6 is already relatively detailed, the Article will benefit from further specification in the form of guidelines and methodologies. Annexes 4A and 4B provide a starting point for such methodologies and guidelines. It will probably not be possible to create one methodology for all types of system. Methods for impact assessments should always be developed and adapted according to technological, legal and societal developments. Adaptation to specific legal, cultural or societal contexts might also be appropriate, especially where the perception of the risks of algorithmic decision-making differ significantly between these contexts.

(6.3.) Article 6(2)(a) requires a description of the relevant features of the system. This description will help experts and the public, but also the implementing authority itself, to understand the functioning of the system and critically evaluate its pros and cons. 'Complexity' was considered as a criterion, but was rejected by the ELI Project Team after consultation with AI experts, who pointed to the lack of definition for such a term. The difficulty of understanding or explaining a system is reflected in the transparency and explainability criteria below.

(6.4.) Article 6(2)(b) addresses the problem that public debates on AI and some other impact assessment models often concentrate on risk alone. This threatens to neglect the potential of ADMSs to improve the work of public administration. But it also aims to avoid exaggerated optimism: some public authorities adopt ADMSs uncritically without assessing their functionality. This might lead to disappointing results and a waste of (financial) resources. Thus, it is important to specify and investigate carefully the potential benefits of the system and the requirements for achieving them (eg system architecture, accessibility for the public, comprehensiveness and quality of the data, etc).

(6.5.) Article 6(2)(c) concerns the external impact of the system. It asks for the investigation of both specific and systemic impacts. Specific impacts are the impacts in the individual use case. Systemic impacts emerge from the fact that previously decentralised decisions are replaced by a few centralised algorithmic systems, which increases the potential damage when these systems fail.[93] Article 6(2)(c)(i) addresses fundamental

rights. The assessment can cover all fundamental rights. The rights that are enumerated are merely those that are most discussed in the context of AI. What rights the individual assessment will need to investigate in depth will depend on the context in which the system is used and should be investigated during the scoping exercise according to Article 5. While the focus of the assessment should be on those rights set out in legally binding documents (eg the EU Charter on Fundamental Rights, the European Convention on Human Rights, or national constitutions), the implementing authority should also consider ethical aspects. Ensuring rights protection beyond the legally required minimum can increase public trust in the system, as public debate often does not differentiate between legal and ethical aspects. Unlike the Commission's Draft Regulation (COM (2021) 206 final), the Project Team did not highlight 'health and safety', as these are most relevant for robotics and automated driving, which are not typical use cases of algorithms in public administration. In contrast to Article 35 GDPR, the assessment also covers societal and environmental aspects, as well as the administrative authority itself. The latter part of the assessment aims to ensure acceptance of the system within the administration and deal with the changes that digitisation inevitably causes.

(6.6.) Article 6(2)(d) asks for the assessment of features and settings of the system in order to manage risks and optimise the system. In contrast to the 'external' impacts of the system in Article 6(2)(c), these are 'internal' features. While Article 6(2)(d)(i) asks what measures are taken to address the external impacts, (ii)–(ix) describe features that any AI system should possess. The meaning of some of the criteria in this paragraph is not completely clear in the legal discussion on AI. In particular, the Project Team discussed the terms 'security' and 'safety'. It understands these terms in accord with IT terminology, where they are not interchangeable. Safety means that the system does not cause external harm (ie it avoids accidents). Security is about preventing unauthorised access to the system and its data. These terms overlap for IT systems as attackers can manipulate the system to cause harm to others.[94] The Project Team also pondered on the exact meaning of, and relationship between, 'transparency' and 'explainability'. Transparency is used here in the sense of making the system's logic visible (eg by publishing the source code), while explainability means that individual decisions can be explained in a meaningful way. Ideally, and as far as technically possible, this means that persons concerned can understand the reasons for individual decisions.

(6.7.) Article 6(2)(e) requires the implementing authority to determine the risk level conclusively, since it only made a preliminary decision on the risk level during the screening procedure.

(6.8.) Article 6(2)(f) requires an overall assessment of the proportionality of the use of the system. The Project Team decided against using the term 'cost-benefit analysis', because this term might imply an overly narrow focus on (economic) effectiveness, while the Project Team would like to encourage a more holistic view that takes societal interests and individual rights into account.

(6.9.) Article 6(2)(g) requires a legality check. Ideally, many issues relating to legality have already been addressed when conducting the assessments required in the previous paragraphs, especially those on fundamental rights. However, the assessor must still check compliance with data protection law or any sectoral legislation. The legality check is the last section of this paragraph, because the Project Team would like the implementing authority to address the assessment with an open mind and consider all impacts of the system before embarking on a comparatively narrow compliance check.

(6.9.) Paragraph 3 refers to the assessment list in Annex 4, which lays out the questions to be answered in even greater detail, but follows the same structure as Article 6. The assessment list is based mainly on the High-Level Expert Group on AI's Assessment List for Trustworthy AI, but has been extended drawing from various other sources. The Project Team prepared both a standard version of the annex (Annex 4A) and an extended version with more detailed questions (Annex 4B). The standard version aims to remain as brief as possible and thus focuses on the questions that the Project Team considered the most important. The extended version can help the implementing authority not to overlook relevant details. The competent legislative bodies have to

[93] Catelijne Muller, 'The Impact of AI on Human Rights, Democracy and the Rule of Law', CAHAI (2020) 06, para 49.
[94] This definition is also implied by the EU High-Level Expert Group's Ethics Guidelines, 16–17.

decide which of these two options they prefer, or whether they prefer to delegate this question to the implementing authority/supervisory authority. It is not mandatory to use the assessment list in every detail. The implementing authority might substitute this list with other, more sector-specific lists.

(6.11.) Paragraph 4 makes it clear that the level of detail of the Article 6 report depends on the individual case and the specific risk levels. It also instructs the implementing authority to aim for a report that is both accurate and understandable. If these aims conflict, the implementing authority must provide a generally understandable summary.

Article 7: Cooperation and Communication with the System Provider and the Data Provider

Comments

(7.1.) Although the main responsibility for conducting the impact assessment and producing the Article 6 report lies with the implementing authority, the reality is that, without active cooperation with the system and data provider, the task would be unmanageable. Additionally, the implementing authority needs to be able to rely on information about the system provided by those who designed, developed, trained and tested it, especially considering that eg descriptions of the development of the system, its technical characteristics and selection of data constitute elements of the Article 6 report.

(7.2.) It is also important to ensure that the information exchange is recorded and reproducible, as mandated in paragraph 1. The main reason for this, other than proving the fulfilment of obligations, is determining the cause and the body liable should the system stop performing as intended due to technical bug, human error, bias in datasets, etc.

(7.3.) To avoid miscommunication and ensure proper operation of the system after its deployment in the public authority, it is crucial to provide a minimum of practical rules of cooperation, such as a joint project team with appointed representatives of the implementing authority, the system provider and the data provider, who have sufficient knowledge and expertise to actively participate in the project team's work.

(7.4.) For the same reasons, paragraph 4 mandates that the system undergoes final testing on the implementing authority's equipment or other equipment used by the implementing authority in the ordinary course of operation (take, for example, cloud services) and this should be done by its employees. The results of such final testing should also be included in the Article 6 report in order to plausibly demonstrate that the system should work as designed when deployed at scale by the implementing authority, because the authority has sufficient equipment and its employees know how to use it.

(7.5.) Finally, paragraph 5 ensures that the purpose of this Article will be accomplished in practice by mandating that its provisions are included in the procurement contract if such a method of acquisition of the system is chosen by the implementing authority. If not – eg if the implementing authority intends to develop the system through its own means – it should nonetheless ensure that the provisions of this Article are taken into account and implemented in the process.

Article 8: Transparency and Protection of Secrets

Sources

Directive (EU) 2019/1024 on open data and the re-use of public sector information; Directive 2009/24/EC on the legal protection of computer programs; Directive 96/9/EC on the legal protection of databases; Council Directive 91/250/EEC on the legal protection of computer programs; Regulation (EU) 2018/1807 on a framework for the free flow of non-personal data in the European Union

Comments

(8.1.) These Model Rules aim at protecting legitimate interests of – on the one hand – the implementing authority and – on the other – the system or data provider, regardless of whether it is a team within the implementing authority, or a separate, private entity that won a tender to develop the system, as well as other individuals and entities which may be involved in the process of conducting impact assessment. Here, in Article 8, the interests that are protected are related to broadly understood 'secrets' – personally identifiable information, privacy, intellectual property, trade secrets (or other legitimate

commercial interests), national security, defence, and public security.

(8.2.) The scope of paragraph 1 is intentionally very broad and non-exhaustive. This is to ensure a consistent level of protection of both individuals' or entities' rights and public interest objectives in various contexts, which inherently leads to overlaps with different legal regimes, for instance, the data protection laws. Since the potential implementation of the Model Rules is not limited to the EU or the Member States, it was crucial to include general clauses related to the types of secrets covered by this paragraph 1. Moreover, this provision applies to both conducting the impact assessment process and drafting its results in the Article 6 report. Moreover, in case of ADMSs marked as high risk systems, to which Article 10 (expert audit and expert board) applies, the experts, whilst conducting the expert audit and drafting the audit report, shall also comply with Article 8 (see paragraph 2).

(8.3.) Publicly sensitive data to be protected and not disclosed at any time during the impact assessment process includes information regarding national security (namely, State security), defence, or public security, in particular where sensitive critical infrastructure protection related information is concerned. Intellectual property and trade secrets of the system are protected, as well as personal data of anyone involved – directly or indirectly – in the process, including where information in an individual training or testing dataset does not present a risk of identifying or singling out a natural person (due to anonymisation or similar process), but when that information is combined with other available information, it could entail such a risk. Similarly, statistical confidentiality should be protected.

(8.4.) For the avoidance of doubt, the term 'intellectual property' refers to copyright and related rights only, including sui generis forms of protection. The term 'trade secrets' covers all forms of protected industrial and commercial property, such as patents, registered trademarks, industrial designs and models, undisclosed know-how and business information.

(8.5.) As to the ADMSs and datasets used for training and testing purposes, Directive 91/250/EEC required all EU Member States to protect computer programs by copyright, as literary works within the meaning of the Berne Convention for the Protection of Literary and Artistic Works. It was codified by Directive 2009/24/EC of the European Parliament and of the Council. Directive 96/9/EC provides for the legal protection of databases, defining a database as 'a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means'. The Directive stipulates that databases are protected both by copyright, which covers intellectual creation, and by sui generis right protecting investment (of money, human resources, effort and energy) in the obtaining, verification or presentation of the content.

(8.6.) Conducting the impact assessment needs to be done using suitable measures to safeguard legally protected secrets. Paragraph 2 establishes a right for the implementing authority and the system or data provider to maintain confidentiality of data related to the impact assessment, limiting this right, however, to duly justified cases. Such a confidentiality restriction should always be weighed against the transparency principle. This is because there is a potential discord between legal protection of broadly understood secrets and transparency of the impact assessment and general principle of transparency of public administration. This means that the protected secrets (whether explicitly listed in paragraph 1 or not) are not of an absolute nature, but rather that they should be always considered in the specific context of a given ADMS and the transparency requirements. This is crucial especially in terms of compliance with the legal protection of computer programs (here: source code of the ADMS) and databases (here: datasets for training and testing purposes), on the one hand, and the notion of explainable AI and public understanding of how the ADMS works, on the other. Therefore, paragraph 3 provides a general framework to enable the assessment of the source code of the ADMS and datasets used for its training and testing (against its reliability, robustness, correctness, adequacy, non-discrimination, etc) ensuring the requirements to protect secrets. The decision regarding limitation or restriction on the access to the source code and the training and testing datasets can be left to the supervisory authority (if such is available). In any case, such a decision needs proper justification.

Article 9: Applicability of this Chapter

(9.) Article 9 clarifies that the Articles on public and expert participation only apply if the system is listed as 'always high risk' in Annex 1, or if the Article 6 report concludes that it poses a high risk. The reason

for this restriction is to avoid excessively burdensome procedures for systems that pose only a relatively modest risk.

Article 10: Expert Audit and Expert Board

Sources

EU High-Level Expert Group on AI, Ethics guidelines for Trustworthy AI 19–20, 36, Rec. 33, Article 5 Environmental Impact Assessment Directive 2011/92/EU; Article 6(3) and Appendix 3 Directive on Automated Decision-Making (Canada); Information Commissioner's Office, AI Auditing Framework (United Kingdom); Opinion of the Data Ethics Commission (Germany) 29, 159–184

Comments

(10.1.) The introduction of the expert audit by the expert board to the impact assessment process (sensu largo) of a high risk system is aimed at providing an extra layer of safeguards before deploying such a system at scale. The added value is in specific technical knowledge and expertise with which the expert board additionally checks the system and searches for threats which might have been missed by the implementing authority. The main subject of the expert audit is the Article 6 report, its completeness, quality, accuracy, etc, but the experts should also have access to the system itself. Such a model of expert audit should give experts the possibility to identify missing elements of the Article 6 report that have not been assessed, and inconsistencies or poor quality, which could indicate that the assessment has not been conducted diligently enough. However, it does not shift the responsibility for the impact assessment from the implementing authority to the experts, and, by not requiring a full audit of the system by the experts, it saves time.

(10.2.) The emphasis in Article 10 is put inter alia on sufficient thematic expertise and independence of the experts, which was deemed crucial for unbiased, reliable audits. Such expertise and independence is ensured by giving the power to appoint and dismiss experts and constitute the expert board by the supervisory authority (and therefore not by the implementing authority, nor the system and data providers) on the basis of objective criteria with considerations for diversity and by introducing rules for avoiding conflicts of interests.

Article 11: Public Participation

Sources

Article 6 Environmental Impact Assessment Directive 2011/92/EU; Article III-25 ReNEUAL Model Rules

Comments

(11.1.) Public consultation in the case of systems with significant risks is a way to ensure that all aspects of their impact have been taken into consideration and to broadly inform the people who will be affected by the system by allowing them to express their views.

(11.2.) Paragraph 1 entrusts the implementing authority with the obligation to carry out the public participation procedure. The implementing authority must ensure that the affected public is able to participate in this procedure. This implies, for example, adapting the methods of providing information about the procedure and the methods of consultation to the public that could be directly or indirectly affected by the ADMS developed.

(11.3.) Information on public consultation is specified in paragraph 2, which gives the minimum elements that must allow the public to know when, on what and by whom it is consulted. At the same time, the supervisory authority is informed of the consultation. Paragraph 3 specifies that the consultation must make available the Article 6 report and the audit report.

(11.4.) Paragraphs 4 and 5 refer to the two modes of consultation in Article III-25 of the ReNEUAL Model Rules. A timeframe for public hearing of two weeks is added in the ReNEUAL Model Rules. At this point it is not proposed to add this extension of the timeframe in order to keep the process simple. The 30 days timeframe could be extended accordingly if needed.

(11.5.) A minimum timeframe of 30 days is given for public consultation (paragraph 6): this timeframe is considered to be sufficient to ensure public participation and not too long to excessively delay the implementation of the ADMS.

(11.6.) In addition to public participation, consultation of other relevant authorities – for instance the data protection authority – is to be undertaken (paragraph 7).

Article 12: Evaluation

Comments

(12.1.) The Article 6 report following the public consultation is completed by the implementing authority. The responses to the audit report by the implementing authority in accordance with Article 10(4) could take place after the public consultation. On the one hand, it allows the implementing authority to include in these responses elements provided by the public, but on the other hand, these responses might not be available to the public.

(12.2.) An extended report is produced by the implementing authority, which includes the initial Article 6 report, the audit report, a summary of the results of the public participation, the evaluation of all contributions (public, expert audit board, other authorities) by the implementing authority and reasoned final opinion on the implementation of the ADMS.

Article 13: Publication

Sources

Article 9(2) Environmental Impact Assessment Directive 2011/92/EU; § 27 Environmental Impact Assessment Act (*Gesetz über die Umweltverträglichkeitsprüfung*) (Germany)

Comments

(13.1.) Paragraph 1 requires publication of the final version of the report (either the Article 6 report or the extended report) at least online. Publication is crucial to ensure transparency of the impact assessment for the public. The implementing authority must keep the report available after the system is out of use, because legal challenges or public criticism of the system might still arise even after it is out of use. This is in line with general laws on file retention, which often mandate quite long periods before files can be destroyed.

(13.2.) Paragraph 2 mandates that the implementing authority give notice of the publication to the experts and members of the public that were consulted under Articles 10 and 11. The implementing authority has a procedural discretion in this regard. It might publish the notice on its website or use an automatic notification system on the consultation website. In addition, the implementing authority must forward the relevant documents to the supervisory authority. This enables the supervisory authority to check whether the report is in conformity with these Model Rules and, if necessary, to exercise its powers under Article 15. It also enables the publication of the report in a central public register (see Article 15 (4)).

(13.3.) Paragraph 3 concerns reports that contain secrets as defined in Article 8. The implementing authority can redact the report to protect these secrets. However, access to the unredacted version is still possible within the limits of the applicable freedom of information rules. The Project Team decided to include this reference because the freedom of information rules establish a broadly accepted balance between the public's interest in information and the protection of secrets.

(13.4.) Paragraph 4 refers to Article 15(4), which governs the publication of the relevant documents in a public register. This could be the public AI register as proposed in Article 60 of the EU Draft AI Regulation (COM (2021) 206 final).

Article 14: Review and Repetition of the Assessment

Sources

Article 35(11) General Data Protection Regulation (EU) 2016/679; Article 43(4) Draft AI Regulation (Commission proposal; COM (2021) 206 final)

Comments

(14.1.) Paragraph 1 requires a review when there is evidence of substantial negative impacts that were not part of the relevant original report (Article 6 report or extended report). Such negative impact can occur because of an unexpected development of self-learning algorithms, because of undetected programming errors, but also because of unexpected behaviour by staff or users. Paragraph 1 also mentions deliberate changes to the system, or the context in which it is used, as a possible cause of substantial negative impacts. Whether an impact is substantial depends on the severity and likelihood of its consequences.

(14.2.) Paragraph 2 requires a review of the report after a certain time, regardless of negative occurrences. Such reviews help to keep the report up to date. They should reflect actual experience in the use of the system. For instance, it will be possible to determine

more accurately the number of false positives/negatives. The periods prescribed are suggestions and might be adapted to reflect practical experiences and the typical speed in which a system changes and/or new experience is gained.

(14.3.) According to paragraph 3, the amended report must again be subject to expert review if this was initially necessary under Article 10. However, experts are free not to make any additional comments. Another public consultation is not prescribed, as it seems excessive to consult the public again, for what might be relatively minor changes. The implementing authority is free to consult the public voluntarily.

Article 15: Supervisory Authority

Sources

Articles 51, 52, 57–59 General Data Protection Regulation (EU) 2016/679; Articles 41, 42, 46, 47–49 Directive (EU) 2016/680 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; Articles 63(5), 64 Draft AI Regulation (Commission proposal; COM (2021) 206 final)

Comments

(15.1.) The impact assessment procedure needs support and oversight by a specialised authority with sufficient resources. As the authority is to supervise other administrative units, it should be independent.

(15.2.) Instead of creating a separate supervisory authority, its tasks and powers could be entrusted to existing authorities that meet the requirements of this paragraph or are adapted to them. In EU Member States, the data protection authorities under the respective EU legislation could be considered for this purpose.

(15.3). Article 15(3) lists the tasks of the supervisory authority. It should support the impact assessment procedure and monitor the implementing authorities' compliance with these rules. In addition, by gathering experience from many assessments and observing developments in the field, it should become a knowledge centre that informs and advises public authorities, affected individuals and organisations and the public on all issues related to the use of systems by public authorities.

(15.4.) The key documents from the various impact assessments should be permanently available to the public on the supervisory authority's website. The register is intended to facilitate the work of public authorities that have to carry out a screening or an impact assessment for the first time, improve future assessments, and promote informed public debate and the responsible use of systems. For secrecy concerns, Article 15(4) refers to Article 13(3).

(15.5.) Article 15(5) enumerates the supervisory powers of the authority. It may, on its own initiative or upon a complaint, investigate whether an impact assessment or repeat impact assessment required by these rules has taken place and whether it has been carried out in accordance with these rules; however, the supervisory authority should not have power to challenge the implementing authority's decision to use a system after a proper assessment, regardless of its result. The supervisory authority should have the right to obtain relevant information notwithstanding the limitations of Article 8, but should maintain the confidentiality set forth in Article 8 in its communications with others. As a response to a missing or unlawful assessment, the supervisory authority can make a (non-binding) recommendation to the implementing authority to stop using the system and, if necessary, to obtain a court order; alternatively, in legal systems where this is compatible with the constitution, the supervising authority could be given the power to issue a binding order. The supervisory authority (and, eventually, the courts) should have some discretion to tolerate minor procedural errors that were immaterial to the outcome of the assessment, or to recommend to correct them rather than discontinue use of the system. In its advisory role, the supervisory authority may make recommendations of any kind, including for systems in lawful use. In addition, the supervisory authority has additional supportive powers under other Articles: to carry out the scoping (where provided for in Article 5(2)), to make available and supervise the experts for the audit under Article 10 and to decide on the access to information under Article 8(3).

Article 16: Complaints and Legal Protection

Sources

Article 9 Aarhus Convention

Comments

(16.1) The rules on legal protection are a minimum standard which, can of course, be strengthened. The rules must be adapted to the specific procedures of each legal system. In particular, in the event of recurrent and unfounded requests, provision can be made for the rapid examination of such complaints.

(16.2.) Paragraph 1 and paragraph 2 put in place the possibility of complaining to the supervisory authority (Article 15). The supervisory authority must react to each complaint. It may reject complaints if they are unsubstantiated, manifestly unfounded or repetitive, and it shall have discretion to adapt the intensity of the investigation to the possible harm of an alleged illegal use of a system. The supervisory authority then could, if it found the complaint well-founded, trigger a procedure to correct the issue (Article 15(5) (c) and (d)) and keep the complainant informed. This mechanism is important in a field which is highly technical and requires special skills: the supervisory authority may be directly alerted by persons affected by the system.

(16.3.) The central issue is to ensure that the persons potentially affected can appeal. In the case of individual decisions, this will generally not be a problem, so that at least the person to whom the decision is addressed can challenge it in court. However, in the case of decisions with a collective scope – for example, let us imagine that the ADMS determines zones related to land use in a city – the persons concerned are not only the owners, but other actors such as NGOs representing collective interests (education, environment, social, etc) could challenge implementation of the system.

(16.4.) Paragraph 3 and paragraph 4 provide that access to court should be conceived in a broad way. However, detailed obligations are not specified, so that these provisions can be implemented in different legal systems. The role of NGOs is stressed for ensuring a collective redress mechanism with access to court following a complaint to the supervisory authority (paragraph 1).

(16.5.) Paragraph 5 recalls some of the characteristics necessary for effective access to justice.

The European Law Institute (ELI) is an independent non-profit organisation established to initiate, conduct and facilitate research, make recommendations and provide practical guidance in the field of European legal development. Building on the wealth of diverse legal traditions, its mission is the quest for better law-making in Europe and the enhancement of European legal integration. By its endeavours, ELI seeks to contribute to the formation of a more vigorous European legal community, integrating the achievements of the various legal cultures, endorsing the value of comparative knowledge, and taking a genuinely pan-European perspective. As such, its work covers all branches of the law: substantive and procedural; private and public.

ELI

EUROPEAN
LAW
INSTITUTE

ISBN: 978-3-9505192-1-1