# ALIGNER D4.2

Methods and guidelines for ethical & law assessment

| Deliverable No. | D4.2 |
|---|---|
| Work Package | WP4 |
| Dissemination Level | PU |
| Author | Donatella Casaburo (KUL) |
| Co-Author | Irina Marsh (CBRNE) |
| Contributor(s) | Plixavra Vogiatzoglou (KUL) |
| Due date | 2023-03-31 |
| Actual submission date | 2023-03-24 |
| Status | Final |
| Revision | 1.0 |
| Reviewed by (if applicable) | Daniel Lückerath (Fraunhofer) |

This document has been prepared in the framework of the European project ALIGNER – Artificial Intelligence Roadmap for Policing and Law Enforcement. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101020574.



**Contact:**

info@aligner-h2020.eu
www.aligner-h2020.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 101020574.

# Executive Summary

The European Commission-funded Coordination and Support Action ALIGNER: Artificial Intelligence Roadmap for Policing and Law Enforcement brings together European actors concerned with Artificial Intelligence, Law Enforcement, and Policing to collectively identify and discuss needs for paving the way for a more secure Europe in which Artificial Intelligence supports law enforcement agencies while simultaneously empowering, benefiting, and protecting the public.

Artificial intelligence can incredibly enhance law enforcement agencies' capabilities to prevent, investigate, detect, and prosecute crimes, as well as to predict and anticipate them. However, despite the numerous promised benefits, the use of AI systems in the law enforcement domain raises numerous ethical and legal concerns. The use made of AI systems by LEAs may not adhere to the four essential ethical imperatives AI practitioners should always strive for: respect for human autonomy; prevention of harm; fairness; and explicability. Moreover, the use of AI systems by LEAs is susceptible to prevent individuals from enjoying some of their fundamental rights, such as: the presumption of innocence and the right to an effective remedy and to a fair trial; the right to equality and non-discrimination; the freedom of expression and information; and the right to respect for private and family life and right to protection of personal data.

ALIGNER's task 4.2 aims to develop a methodological approach to adequately address the critical ethical and legal concerns related to the use of AI systems in the law enforcement domain. The outcome of task 4.2 is deliverable D4.2 – Methods and guidelines for ethical & law assessment. The deliverable provides a fundamental rights impact assessment template, suitable to be integrated in the governance systems of LEAs planning to deploy AI systems for law enforcement purposes.

# Table of Contents

# List of Abbreviations

| Abbreviation | Meaning |
|---|---|
| AFRIA | ALIGNER Fundamental Rights Impact Assessment |
| AI | Artificial Intelligence |
| CFREU | Charter of Fundamental Rights of the European Union |
| DPA | Data protection supervisory authority |
| DPIA | Data protection impact assessment |
| EIA | Ethical impact assessment |
| EU | European Union |
| FRIA | Fundamental Rights Impact Assessment |
| GDPR | 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) |
| LEA | Law Enforcement Agency |
| LEAAB | Law Enforcement Agency Advisory Board |
| LED | Directive 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, and repealing Council Framework Decision 2008/977/JHA (Law Enforcement Directive) |
| SIEAB | Scientific, Industrial and Ethical Advisory Board |
| WP | Work package |

# 1. Introduction

The H2020 project ALIGNER brings together European practitioners from law enforcement and policing, civil society, policymaking, research, and industry with the objective of discussing opportunities, challenges, needs, and risks emerging from the use of artificial intelligence (AI) technologies in the law enforcement field. The outcome of these discussions will be systematised in a research and policy roadmap, meeting the operational, cooperative, and collaborative needs of law enforcement agencies.
ALIGNER's work package (WP) 4 – Ethics & Law aims to map the current ethics and law landscape concerning artificial intelligence solutions for law enforcement, in order to set up and maintain a systematic assessment process for (novel) artificial intelligence solutions with potential for enhancement of the law enforcement agencies' work.

Artificial intelligence can incredibly enhance law enforcement agencies' (LEAs) capabilities to prevent, investigate, detect, and prosecute crimes, as well as to predict and anticipate them. In the European Union (EU), LEAs are already deploying AI systems in their daily operations, for instance, for: patrolling hazardous areas; gathering and analysing data collected in a crime scene to obtain relevant evidence; identifying items, crime suspects, or victims; or forecasting individuals or geographical areas with an increased probability of criminal activity.

Despite the numerous promised benefits, the use of AI systems in the law enforcement domain raises numerous ethical and legal concerns. The use made of AI systems by LEAs may not adhere to the four, essential, ethical imperatives AI practitioners should always strive for. These four ethical principles are: respect for human autonomy; prevention of harm; fairness; and explicability.[1]

Moreover, the use of AI systems by LEAs is susceptible to prevent individuals from enjoying their fundamental rights. In particular, individuals may suffer significant interferences in the exercise of the following protected rights: presumption of innocence and right to an effective remedy and to a fair trial; right to equality and non-discrimination; freedom of expression and information; and right to respect for private and family life and right to protection of personal data.[2]

ALIGNER's task 4.2 aims to develop a methodological approach to adequately address the critical ethical and legal concerns related to the use of AI systems in the law enforcement domain. The outcome of task 4.2 is deliverable D4.2 – Methods and guidelines for ethical & law assessment. The deliverable provides a fundamental rights impact assessment template, suitable to be integrated in the governance systems of LEAs planning to deploy AI systems for law enforcement purposes.

---

[1] According to the High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, (accessed on 26 February 2023), these four ethical imperatives are translated into the following seven key requirements that an AI system should implement: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental wellbeing; and accountability.
[2] E. Eren, D. Casaburo and P. Vogiatzoglou, 'ALIGNER D4.1 – State-of-the-art reports on ethics & law aspects in Law Enforcement and Artificial Intelligence', 2022.

## 1.1 Gender Statement

The deliverables of WP4 aim to support a wide array of stakeholders (public, private, and third sector). Partners of WP4 ensured – to the best of their capacities – a balanced gender representation during the co-development of the ALIGNER Fundamental Rights Impact Assessment templates with the advisory boards and experts.

This deliverable was gender-proofed during the internal review process under and in accordance with a gender-proofing checklist described in Deliverable 1.2 “Project Handbook”.

## 1.2 Relation to other deliverables

ALIGNER D4.1 mapped out the ethical and legal framework applicable to AI systems used for law enforcement purposes. Thus, D4.1 laid the groundwork for this deliverable: the ethical norms and fundamental rights illustrated therein are used as benchmarks for constructing a fundamental rights impact assessment template addressed to LEAs aiming to deploy AI systems.

The fundamental right impact assessment template proposed in this deliverable complements the AI technology impact assessment framework, which will be substantiated in ALIGNER D3.2.

## 1.3 Structure of this report

This document consists of four sections. The first section [§ 2.] gives an overview of the main existing methodologies for conducting ethical and legal assessments of information technologies, as well as of the principal issues related to their application regarding AI systems used in the law enforcement domain. The second section [§ 3.] presents the ALIGNER Fundamental Rights Impact Assessment and its methodology, while the third section [§ 4.] contains the template itself. Finally, the fourth and last section [§ 5.] illustrates the efforts undertaken to validate both the ALIGNER Fundamental Rights Impact Assessment methodology and template.

# 2. State-of-the-art in ethical and law assessments

As explained above [§ 1.], LEAs can gain enormous advantages from the use of AI systems. However, the same AI systems can create numerous ethical and legal concerns, especially when used in the sensitive domain of law enforcement. To ensure the trustworthiness of the deployed AI systems, LEAs need to adequately assess and address these sets of concerns, from both an ethical and a legal perspective.

Throughout the years, both organisations and policy-makers have made use of impact assessment frameworks to identify problems and possible solutions related to a planned activity. In the field of information technologies, ethical impact assessments and data protection impact assessments are of particular relevance. While the first assessment can be performed on a voluntary basis, the second one is often a legal obligation.

In the EU, there is not a single and institutionalised methodology to conduct an ethical and data protection impact assessment of AI systems. However, many national authorities, public and private organisations, and scholars have developed non-binding guidelines, methodologies, or templates. The available contributions need to be mapped out [§ 2.1] to evaluate whether any of them is suitable to be used by LEAs to adequately address the ethical and legal concerns related to law enforcement AI [§ 2.2].

## 2.1 Ethical and data protection impact assessments

The following paragraphs contain an overview of the main important contributions in the field of ethics [§ 2.1.1] and data protection [§ 2.1.2], including the ethical and legal assessment methodology used during the H2020 project MAGNETO[3] [§ 2.1.3].

### 2.1.1 Ethical impact assessment

An ethical impact assessment (EIA) can be defined as a "*process during which an organisation, together with stakeholders, considers the ethical issues or impacts posed by a new project, technology, service, program, legislation, or other initiative, to identify risks and solutions*".[4] EIAs are often presented as instruments complementary to data protection impact assessments [§ 2.1.2], aimed to ensure an adequate analysis of the ethical implications of a project.[5]

EIAs have two main characteristics.[6] First, they are only focused on ethical impacts, namely outcomes that have ethical relevance or that raise ethical issues. Second, they do not merely observe impacts, but they also evaluate them and their acceptability from an ethics point of view.

---

[3] MAGNETO, 'Fighting Against Crime and Terrorism', http://magneto-h2020.eu/, (accessed on 22 February 2023).
[4] D. Wright, 'Ethical Impact Assessment', in J. Britt Holbrook and C. Mitcham (eds.), *Ethics, Science, Technology and Engineering: A Global Resource*, Macmillan Reference USA, 2014, as cited in W. Reijers *et al*., 'SATORI D4.1 – Annex 1: A Common Framework for Ethical Impact Assessment', 2016, https://satoriproject.eu/media/D4.1_Annex_1_EIA_Proposal.pdf, (accessed on 22 February 2023).
[5] D. Wright and E. Mordini, 'Privacy and Ethical Impact Assessment', in D. Wright and P. De Hert, *Privacy Impact Assessment*, Springer, 2012, pp. 397-418.
[6] C. Shelley-Egan *et al*., 'SATORI D1.1 – Ethical Assessment of Research and Innovation: A Comparative Analysis of Practices and Institutions in the EU and selected other countries', 2016, https://satoriproject.eu/media/D1.1_Ethical-assessment-of-RI_a-comparative-analysis-1.pdf, (accessed on 22 February 2023), p. 30.

Scholars and organisations have developed various EIA methodologies. The FP7 project SATORI (Stakeholders Acting Together on the Ethical Impact Assessment of Research and Innovation) has conducted an extensive comparative analysis of the existing EIA approaches and practices across different fields, in both EU and non-EU countries.[7] The results of this analysis are summarized in a report and 47 annexes, each of them focused on a different topical issue.[8] For instance, Annex 2.b.1 is dedicated to the EIA of information technologies.[9] In the latter, the authors highlight how, despite a general agreement on the central ethical concerns related to information technologies,[10] there is not an institutionalised and sector-specific EIA methodology.[11]

Among the various methodologies proposed by scholars to assess and address the ethical impact of novel technologies, of particular relevance are the EIA framework outlined by D. Wright,[12] and the DIODE meta-methodology of I. Harris *et al*.[13]

**FRAMEWORK FOR EIA OF INFORMATION TECHNOLOGY** – The framework can be used for performing an EIA of any policy or project involving information technologies, in order to ensure that their ethical implications are adequately examined and mitigation measures are taken as necessary.

In the framework, the values that must be pursued while deploying information technologies are clustered together into five overarching ethical principles. The first four principles were originally posited by Beauchamp and Childress;[14] the fifth is added by the author. The five benchmarks of the framework are:

1. Respect for autonomy, i.e., the right to liberty. This entails: the value of dignity; and the need for an informed consent.
2. Nonmaleficence, i.e., the obligation to avoid harm. This entails: the need for ensuring safety; the realisation of social solidarity; the need to avoid isolation; and the prohibition of discrimination.
3. Beneficence, i.e., the obligation to provide benefits and to produce the best overall results. This entails: the obligation to provide universal service; the need for accessibility; value sensitive design; and the need to ensure the sustainability of the technology.
4. Justice, i.e., the right to a fair, equitable and appropriate treatment. This entails the need to ensure both equality and fairness.

---

[7] C. Shelley-Egan *et al*., 'SATORI D1.1 – Ethical Assessment of Research and Innovation: A Comparative Analysis of Practices and Institutions in the EU and selected other countries', https://satoriproject.eu/media/D1.1_Ethical-assessment-of-RI_a-comparative-analysis-1.pdf, (accessed on 22 February 2023).
[8] SATORI, 'Work Package 1: Comparative Analysis of Ethics Assessment Practices – Deliverable 1.1: Ethical Assessment of R&I: A Comparative Analysis', https://satoriproject.eu/work_packages/comparative-analysis-of-ethics-assessment-practices/, (accessed on 22 February 2023).
[9] J. Hartz Sørake and P. Brey, 'SATORI D1.1 – Annex 2.b.1: Ethics Assessment in Different Field – Information Technologies, 2015, https://satoriproject.eu/media/2.b.1-Information-technology.pdf, (accessed on 22 February 2023).
[10] According to J. Hartz Sørake and P. Brey, 'SATORI D1.1 – Annex 2.b.1: Ethics Assessment in Different Field – Information Technologies, https://satoriproject.eu/media/2.b.1-Information-technology.pdf, (accessed on 22 February 2023), these are: privacy; security and crime; free expression and content control; equity and access; intellectual property; IT and responsibility; autonomy, sociality, and authenticity; AI and robotics; and embedded values.
[11] J. Hartz Sørake and P. Brey, 'SATORI D1.1 – Annex 2.b.1: Ethics Assessment in Different Field – Information Technologies, https://satoriproject.eu/media/2.b.1-Information-technology.pdf, (accessed on 22 February 2023), p. 13.
[12] D. Wright, 'A framework for the ethical impact assessment of information technology', *Ethics and Information Technology*, vol. 13, no. 3, 2011, pp.199-226.
[13] I. Harris *et al*., 'Ethical assessment of new technologies: a meta-methodology', *Journal of Information, Communication & Ethics in Society*, vol. 9, no. 1, 2011, pp. 49-64.
[14] T. L. Beauchamp and J. F. Childress, *Principles of Biomedical Ethics*, Oxford University Press USA, 2013.

5. Privacy and data protection, i.e., privacy of the person, personal behaviour, personal communications and personal data. This entails the principles of: data minimisation; data quality; purpose specification; use limitation; confidentiality, security and protection of data; transparency; access; anonymity.

The framework includes questions for each of the overarching principles and related values, so to facilitate the consideration of the ethical issues that may arise.

The EIA framework is also supported by a catalogue of tools (e.g., surveys, checklists, ethical matrixes) and procedural practices (e.g., stakeholder consultations, risk assessments, audits) that can help decision-makers both understand how the technology is ethically perceived by stakeholders and ensure the transparency of the assessment process.

**DIODE META-METHODOLOGY** – The DIODE meta-methodology can be used by business decision-makers, technologists, and inventors for the ethical assessment of new and emerging technologies.

The meta-methodology is based on the fundamental ethical principles revealed by the Charter of Fundamental Rights of the European Union (CFREU).[15] These ethical principles are:

1. Rights of individuals;
2. Educational rights and freedoms;
3. Non-discrimination rights;
4. Environmental concerns; and
5. Justice

The meta-methodology aims to have an appropriate and manageable scope, which takes into account three different perspectives: the governmental, the organisational and the individuals' ethics. Moreover, it is suitable to evaluate and mitigate the ethical impact of both abstract technologies and specific applications.

The DIODE meta-methodology consists of five stages:

1. Define questions. Identification of the technology or project to be examined, in order to be able to frame the ethical questions.
2. Issues analysis. Consideration of all the possible affected parties and exam of the risks and rewards.
3. Options evaluation. Assessment of relevant choices, as well as of the appropriate safeguards.
4. Decision determination. Statement of the ethical decisions made and their motivation, including the circumstances that may lead to a revision of the decision.
5. Explanations dissemination. Appropriate communication of the decisions made.

[15] Charter of Fundamental Rights of the European Union, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012P%2FTXT, (accessed on 22 February 2023).

### 2.1.2 Data Protection Impact Assessment

As already explained in ALIGNER D4.1 – State-of-the-art reports on ethics & law aspects in Law Enforcement and Artificial Intelligence,[16] the Law Enforcement Directive[17] (LED) governs the processing of personal data in the law enforcement context.[18] Article 27 LED foresees an important measure to help LEAs acting as personal data controllers ensure and demonstrate compliance with the data protection legislation, i.e., the data protection impact assessment (DPIA).

A DPIA is "*a process designed to describe the processing* [...] *and help manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data by assessing them and determining the measures to address them*".[19]

In line with the risk-based approach adopted by the LED,[20] LEAs are not obliged to perform a DPIA for all the personal data processing operations they undertake. However, performing a DPIA becomes mandatory when the processing – especially if performed through the means of new technologies – is likely to result in a high risk to the rights and freedoms of natural persons.[21] Here, the reference to "*the rights and freedoms of natural persons*" implies that, while assessing the potential risks caused by the processing of personal data, attention should be paid not only to the right to data protection but also to other fundamental rights which may equally be affected.

A DPIA should be performed for the first time "*prior to the processing*",[22] so to help LEAs in deciding *whether* and *how* to conduct the personal data processing operations. However, performing a DPIA is not a one-time exercise, but a continuous process: especially when the processing is subject to changes, the DPIA should be continuously reviewed and re-assessed.[23]

Article 27 LED briefly prescribes the minimum requirements of a DPIA:

1. Description of the processing operations. LEAs should describe the envisaged personal data processing operations in general, taking into account the nature, scope, context, and purposes of the processing.
2. Identification and assessment of risks. LEAs should assess the (likelihood and severity of the) potential risks to the rights and freedoms of data subjects associated to the envisaged processing operations.

---

[16] E. Eren, D. Casaburo and P. Vogiatzoglou, 'ALIGNER D4.1 – State-of-the-art reports on ethics & law aspects in Law Enforcement and Artificial Intelligence', p. 42.
[17] 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, and repealing Council Framework Decision 2008/977/JHA, LED, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680, (accessed on 20 February 2023).
[18] LED, Article 1, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680, (accessed on 20 February 2023).
[19] Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', 2017, https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711, (accessed on 20 February 2023), p. 4.
[20] LED, Article 19, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680, (accessed on 20 February 2023).
[21] LED, Article 27, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680, (accessed on 20 February 2023).
[22] *Ibid*.
[23] Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711, (accessed on 20 February 2023), p. 14.

3. Mitigation measures. LEAs should identify potential mitigation measures to address and reduce the identified risks.
4. Security measures, safeguards and mechanisms. LEAs should identify additional measures suitable to ensure and demonstrate compliance with the data protection legislation.
5. Documentation. LEAs should document and maintain records of the whole DPIA process.

The LED does not prescribe a mandatory template for performing a DPIA, but leaves LEAs some flexibility in determining its structure and form.[24]

In the last years, various different DPIA guidelines and methodologies have been developed by (national) data protection supervisory authorities (DPAs), scholars or practitioners. However, the vast majority of the contributions is primarily addressed to (private or public) organizations acting for non-law enforcement purposes and is, thus, based on the data protection requirements foreseen in the General Data Protection Regulation (GDPR).[25]

**NON-LAW ENFORCEMENT-RELATED GUIDANCE** – At a European level, in 2017, the Article 29 Data Protection Working Party published a set of guidelines on DPIAs, including a brief checklist of the criteria for an acceptable DPIA.[26] This can be used by personal data controllers to assess whether their own DPIA methodology is sufficiently comprehensive to ensure full GDPR compliance.

At a national level, several DPAs have provided more detailed guidelines on what a DPIA should assess and how this assessment should be performed. For instance, this is the case of the Maltese Information and Data Protection Commissioner[27] and of the Spanish Agencia Española Protección Datos, whose guidelines are differentiated on the basis of the public[28] or private[29] nature of the organisation.

Additionally, other DPAs have drafted DPIA templates, ready to be immediately implemented in an organisation's governance systems. The French Commission Nationale de l'Informatique et des Libertés has produced a Privacy Impact Assessment toolkit, consisting of three guides (methodology,

[24] Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711, (accessed on 20 February 2023), p. 17.
[25] 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), GDPR, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679, (accessed on 20 February 2023).
[26] Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711, (accessed on 20 February 2023), p. 22.
[27] Information and Data Protection Commissioner, 'Guidelines on DPIA template', https://idpc.org.mt/wp-content/uploads/2020/07/Guidelines-on-DPIA-template.pdf, (accessed on 20 February 2023).
[28] Agencia Española Protección Datos, 'Template for Data Protection Impact Assessment Report (DPIA) For Public Adminisitrations', 2022, https://www.aepd.es/es/documento/modelo-informe-EIPD-AAPP-en.rtf, (accessed on 20 February 2023).
[29] Agencia Española Protección Datos, 'Template for Data Protection Impact Assessment Report (DPIA) For Private Sector', 2022, https://www.aepd.es/es/documento/modelo-informe-EIPD-sector-privado-en.rtf, (accessed on 20 February 2023).

templates, and knowledge bases) and a software.[30] Similarly, the UK Information Commissioner's Office released a sample DPIA template,[31] along with a detailed guide.[32]

Among the various legal scholars' contributions on the topic, of particular relevance is that given by Kloza *et al*.[33] In their policy brief, the authors propose an extensive DPIA template based on a critical and comparative analysis of (some of) the already existing technical standards and methodologies.

**LAW ENFORCEMENT-RELATED GUIDANCE** – With regard to DPIAs in the law enforcement domain, the number of existing contributions is definitely more scarce, and also of a lower level of detail. To date, there are no sector-specific methodologies or templates for performing a DPIA, but only broader guidelines issued by DPAs and legal scholars.

At a European level, neither the Article 29 Data Protection Working Party nor the European Data Protection Board have drafted LED-specific guidelines – contrarily to what was done for the GDPR. However, at a national level, relevant resources are coming from the UK and Slovenia.

The UK Information Commissioner's Office drafted its guide to law enforcement processing.[34] The guide's section on DPIAs is limited to a brief explanation of Article 27 LED, while it redirects to the corresponding GDPR guide for more detailed guidelines.[35] The more general section of the guide on accountability and governance refers to the 'Toolkit for organisations considering using data analytics', an online tool developed by the same UK DPA.[36] The toolkit aims to help organizations planning to deploy data analytics techniques (including AI systems) consider risks, rights and freedoms of individuals in the context of data protection law. Based on the answers given to a series of questions, the toolkit produces a final report containing tailored advice on how to improve data protection compliance. While the toolkit does not substitute a DPIA, it can help LEAs establish their own DPIA methodology.

In 2014, the Information Commissioner of the Republic of Slovenia published some privacy impact assessment guidelines for the introduction of new police powers.[37] The purpose of the document is to guide law enforcement policy-makers in their decisions on the establishment of new police powers,

---

[30] Commission Nationale de l'Informatique et des Libertés, 'Privacy Impact Assessment (PIA)', https://www.cnil.fr/en/privacy-impact-assessment-pia, (accessed on 20 February 2023).
[31] Information Commissioner's Office, 'Sample DPIA template', June 2018, https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fico.org.uk%2Fmedia%2Ffor-organisations%2Fdocuments%2F2553993%2Fdpia-template.docx&wdOrigin=BROWSELINK, (accessed on 20 February 2023).
[32] Information Commissioner's Office, 'Data Protection Impact Assessments (DPIAs)', https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/, (accessed on 20 February 2023).
[33] D. Kloza *et al*., 'Data protection impact assessment in the European Union: developing a template for a report from the assessment process', *d.pia.lab Policy Brief No. 1/2020*, 2020, https://cris.vub.be/ws/portalfiles/portal/53602836/dpialab_pb2020_1_final.pdf, (accessed on 20 February 2023).
[34] Information Commissioner's Office, 'Guide to Law Enforcement Processing', https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-le-processing/, (accessed on 20 February 2023).
[35] Information Commissioner's Office, 'Data Protection Impact Assessments (DPIAs)', https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/, (accessed on 20 February 2023).
[36] Information Commissioner's Office, 'Toolkit for organizations considering using data analytics', https://ico.org.uk/for-organisations/toolkit-for-organisations-considering-using-data-analytics/, (accessed on 20 February 2023).
[37] Information Commissioner of the Republic of Slovenia, 'Privacy Impact Assessment (PIA) Guidelines for the Introduction of new Police Powers', 2014 https://www.ip-rs.si/fileadmin/user_upload/Pdf/smernice/PIA_guideliness_for_introduction_of_new_police_powers_english.pdf, (accessed on 20 February 2023).

especially when these entail technological tools. Therefore, the main addressees of the guidelines are law- and policy- makers, and not LEAs responsible for ensuring data protection compliance. Nonetheless, LEAs can benefit from the methodology outlined in the DPA guidelines to establish their own DPIA methodology.

Among the legal scholars, Marquenie and Quezada-Tavárez highlighted the importance of performing a DPIA for novel data-driven applications used for law enforcement purposes.[38] The authors identified the most common legal and ethical concerns related to the use of AI systems in a law enforcement context; these include, for instance: fundamental rights protection; unfairness and opacity; and police integrity. DPIAs should adequately assess all these concerns and foresee appropriate technical and organisational mitigation measures.

### 2.1.3 The MAGNETO methodology

The H2020 project MAGNETO[39] (Multimedia Analysis and correlation engine for organised crime prevention and investigation) aimed to develop a platform to support LEAs to fuse and analyse multiple data sources through analysing massive volumes of heterogenous and fragmented data for the prevention, investigation and prosecution of criminal offences.[40] In order to mitigate the impact of MAGNETO on ethical values, data protection and other fundamental rights, the consortium worked together to implement a comprehensive ethical and legal impact assessment framework.

MAGNETO's ethical and legal impact assessment methodology was developed following the risk-based and fundamental rights compliant approach of the regulatory changes in the EU. In this context, the notion of 'risk' is to be regarded as the not-so-improbable likelihood of a negative event. A risk-based approach entails acting in a proactive rather than a reactive manner to address the concerns that may be related to an activity. MAGNETO's methodology underscores the proactive role taken by the consortium when taking the necessary steps to minimise the risks to the rights and freedoms of the individuals that may be impacted by the MAGNETO system.

Following the development and the implementation cycles of the MAGNETO system, MAGNETO's ethical and legal impact assessment methodology is three tiered and consists of: a Legal and Ethical Checklist,[41] a DPIA[42] and an EIA.[43]

**LEGAL AND ETHICAL CHECKLIST** – The Checklist encompasses a hands-on practice of the risk-based approach. It provides an overview of the key legal requirements and translates them into technical strategies for the implementation in the system. The checklist is primarily addressed to the technical partners, to be consulted and updated at different stages of development of the MAGNETO platform

[38] T. Marquenie and K. Quezada-Tavárez, 'Data Protection Impact Assessments in Law Enforcement', in G. Markarian *et al*. (eds.), *Security Technologies and Social Implications*, Wiley-IEEE Press, 2023, pp. 32-60.
[39] MAGNETO, 'Fighting Against Crime and Terrorism', http://magneto-h2020.eu/, (accessed on 25 February 2023).
[40] T. Marquenie *et al*., 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', 2019, https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 25 February 2023).
[41] T. Marquenie *et al*., 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 25 February 2023), pp. 26-27.
[42] T. Marquenie *et al*., 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 25 February 2023), pp. 17-26.
[43] T. Marquenie *et al*., 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 25 February 2023), pp. 27-38.

and tools. Two methodological approaches were merged to provide specific advice on value sensitive design for each of the MAGNETO tools: the anticipatory technology ethics and the fundamental rights impact assessment.

**EIA** – The EIA consists of two forms: the Ethical Risk Assessment Form and the Misuse Risk Assessment Form.

The Ethical Risk Assessment Form explores the extent to which each MAGNETO tool and the integrated MAGNETO system challenge the established ethical values. The Ethical Risk Assessment Form draws from the High-Level Expert Group's Ethics Guidelines for Trustworthy AI,[44] adapted to the particularities of AI surveillance systems. The Ethical Risk Assessment Form has a set of nine key requirements[45] the MAGNETO system should meet to be deemed trustworthy. Each of requirements includes specific challenges and asks for control actions to mitigate the risk and to reduce it to 'as low as reasonably practicable'.

The Misuse Risk Assessment Form engages in anticipatory risk assessment in respect to the possibility of misuse of each of the MAGNETO tool and the integrated MAGNETO system. The Misuse Risk Assessment Form builds on the three moral risks related with surveillance technologies identified in the FP7 project SURVEILLE (Surveillance: ethical issues, legal limitations and efficiency).[46] These are: the moral risk of privacy intrusion; the moral risk of error; and the moral risk of damage to relations of trust. For each of the three moral risks, the Form identifies specific challenges and asks for control actions to mitigate the risk and reduce it to 'as low as reasonably practicable'.

Both the Ethical Risk Assessment and the Misuse Risk Assessment forms are methodologically based on a risk assessment approach which allows for an integrated cycle that comprises the identification, assessment, and prioritization of ethical risks, followed by the coordinated and efficient use of resources to monitor, minimize, and control the probability and/or impact of the risks occurring.[47] Both forms use the same fundamental model for the calculation of risk, the Ethical Risk Calculation System:[48] according to it, the ethical risk can be calculated from an assessment of the probability (P) of an event occurring and the severity (S) of the consequences if it does occur [i.e., ER = P * S]. The results obtained are for guidance and for further discussion only.

**DPIA** – The DPIA methodology identifies the data protection risks associated with the personal data processing activities carried out within the MAGNETO system, as well as the measures for managing these risks. The DPIA methodology serves as a starting point to allow LEAs to perform the DPIAs potentially required by the national laws prior to the use of MAGNETO.

---

[44] High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, (accessed on 25 February 2023),

[45] The nine key requirements are: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental wellbeing; accountability; respect of legal restriction on the development, use and export of the technology; and provision of training for users.

[46] J. Guelke, 'SURVEILLE D2.2 – Paper with input from end-users', 2013, https://surveille.eui.eu/wp-content/uploads/sites/19/2015/04/D2.2.Paper-with-Input-from-End-Users.pdf, (accessed on 25 February 2023).

[47] I. Marsh, N. Hale, and D. Kelly, *Ethical Assessment Regarding the Use or Misuse of AI Systems for Law Enforcement. A Handbook for Law Enforcement Officials*, 2021, pp 20-21.

[48] T. Marquenie *et al*., 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 25 February 2023), pp. 34-35.

## 2.2 Gap analysis

As illustrated above [§§ 2.1.1. and 2.1.2], there are many frameworks, methodologies, and templates for performing EIAs and DPIAs. However, none of the already existing frameworks is suitable to fully address the ethical and legal concerns related to the use of AI systems in a law enforcement context [§ 1.].

**GAPS IN EXISTING EIA METHODOLOGIES** – The existing information technologies-specific EIA methodologies have two main functions, i.e., providing practitioners with a list of ethical concerns that need to be addressed and illustrating the necessary steps to do so. It is true that any proposed EIA framework or methodology that aims to be deployed for more than one projects needs to bear a certain degree of genericity and, thus, requires some adaptation efforts. However, the existing EIA methodologies completely lack operationality: they do not require adaptation, but rather concretisation into a tool that can be integrated in the LEAs' governance systems (e.g., a checklist or a template).

The proposed EIA methodologies are based on the ethical principles established for information technologies. AI systems differ from other, even sophisticated, information technologies, due to their capability to: interpret data; reason on the knowledge derived from them; and make independent decisions to achieve a certain goal. The peculiarities of AI systems raise particular ethical concerns, which are even stronger when the AI systems are deployed for law enforcement purposes.[49] To address these concerns, within the broader field of the ethics of technology, a new 'ethics of AI' was developed. Its ethical principles have to be used as benchmarks for any EIA tool primarily focused on the assessment of AI systems.

**GAPS IN EXISTING DPIA METHODOLOGIES** – The existing law enforcement-related guidance on DPIAs still needs to be translated into an operative tool, suitable to be immediately integrated in the governance systems of LEAs. Despite that, LEAs can rely on the GDPR-based DPIA templates drafted by national DPAs and practitioners, provided that they implement some (minor) adaptations, e.g., for what concerns the transparency obligations and compliance with data subjects' rights.

As such, the existing DPIA templates can – and must – be used by LEAs to help manage the risks to the right to data protection of individuals that are resulting from the processing of personal data, also when the processing is performed through the means of an AI system. However, and contrarily to what is required by the LED,[50] the proposed DPIA templates fail to adopt a broader and more comprehensive fundamental rights perspective. This is particularly problematic in the law enforcement domain, where the fundamental rights at stake are of outmost importance and the risk of abuse high. Therefore, any DPIA template aimed to be used by LEAs needs to be further expanded, so to also address the other relevant fundamental rights concerns.

[49] Examples of relevant concerns may be related to the possible unfairness and opacity of AI's outcomes, the difficulties in establishing accountability for the mistakes committed or the harm caused, the reduced room for human intervention in the decision-making process.
[50] LED, Article 27, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680, (accessed on 24 February 2023).

**GAPS IN THE MAGNETO METHODOLOGY** – MAGNETO's Legal and Ethical Checklist[51] is aimed to be filled by technical developers during the design and development stage of the technology. Therefore, it is not suitable to be used by LEAs during the deployment phase.

MAGNETO's Ethical Risk Assessment Form[52] fills some of the gaps identified in the other existing EIA methodologies. First, the Form is an operational tool, i.e., a template suitable to be immediately implemented in LEAs' governance systems. Second, the Form is based on the High-Level Expert Group's Ethics Guidelines for Trustworthy AI[53] and, therefore, uses as benchmarks the ethical principles of the ethics of AI. However, the Form does not allow LEAs to separately evaluate each of the ethical challenges identified, and this can lead to less accurate results.

MAGNETO's Misuse Risk Assessment Form aims to address the possible risks related to the use of surveillance technologies.[54] Therefore, it is only relevant for AI systems used for surveillance purposes. However, even for this subset of AI systems, the Form is of questionable added value, for two main reasons. First, the data protection-related risks are already adequately addressed via the mandatory DPIA. Second, the Form is too broad and generic to lead to a sufficiently accurate and comprehensive assessment of all the possible fundamental rights concerns.

MAGNETO's DPIA methodology offers guidance on how to perform a DPIA for AI systems deployed in a law enforcement context.[55] The methodology promotes the adoption of a risk-based approach regarding all fundamental rights. Accordingly, the risk assessment of the MAGNETO systems and tools addresses, for instance: interferences with due process guarantees; interferences with the right to privacy; discrimination and biases.[56] Despite the non-exhaustivity of the list of fundamental rights considered, this approach represents a good attempt to broaden the scope of DPIAs to transform them into an instrument suitable to "*manage the risks to the rights and freedoms of natural persons*".[57] Unfortunately, the methodology is not further concretised into an operational tool for LEAs.

The state-of-the-art analysis on EIAs and DPIAs for AI systems used for law enforcement purposes reveals a research gap that needs to be filled. Despite the existence of many frameworks, methodologies, and templates in both fields, LEAs do not have at their disposal an instrument that:

---

[51] T. Marquenie *et al.*, 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 27 February 2023), pp. 26-27.
[52] T. Marquenie *et al.*, 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 24 February 2023), pp. 30-36.
[53] High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, (accessed on 24 February 2023).
[54] T. Marquenie *et al.*, 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 24 February 2023), pp. 36-38.
[55] T. Marquenie *et al.*, 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 24 February 2023), pp. 17-27.
[56] T. Marquenie *et al.*, 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 24 February 2023), p. 92.
[57] Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711, (accessed on 24 February 2023), p. 4.

- Is operational, i.e., suitable to be directly implemented in the governance system of the organisation;
- Addresses the ethical concerns raised by the AI ethics; and
- Addresses the fundamental rights concerns.

# 3. The ALIGNER Fundamental Rights Impact Assessment

As concluded above [§ 2.2], to fully address the critical ethical and fundamental rights concerns related to the use of AI systems for law enforcement purposes [§ 1.], LEAs need a new operational instrument.

Such an operational instrument should be composed by two connected and complementary templates. On the one hand, a first template aimed to address the ethical concerns can be built upon the ethical and legal assessment methodology drafted for the H2020 project MAGNETO, and especially upon the Ethical Risk Assessment Form. This needs to be revised to better reflect the perspective of LEAs; updated to encompass the latest developments in the AI legislation and policy; and simplified [§ 3.1]. On the other hand, a second template aimed to address the fundamental rights concerns should be created based on existing methodologies [§ 3.2].

The methodology of the ALIGNER Fundamental Rights Impact Assessment is 19xplainned in the last paragraph of the current section of this document [§ 3.3].

## 3.1 Revising MAGNETO's Ethical Risk Assessment Form

To fully address the ethical concerns related to the use of AI systems done by LEAs, MAGNETO's Ethical Risk Assessment Form[58] needs to be further targeted, updated, and simplified.

**TARGET** – The template is directed at LEAs aiming to deploy AI systems in their daily operations. Therefore, it needs to address the ethical concerns related to the deployment stage of the technology, and not those instead related to the research and development stage.

**UPDATE** – Where necessary, the ethical principles established in the High-Level Expert Group's Ethics Guidelines for Trustworthy AI[59] need to be updated, to reflect the latest developments in the AI policy and legislation. To do so, two main kinds of sources need to be consulted: the most recent AI assessment frameworks developed by public and private organisations[60] and the European Commission's proposal for an AI Act.[61]

---

[58] T. Marquenie *et al.*, 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 27 February 2023), pp. 30-36.

[59] High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, (accessed on 27 February 2023).

[60] All the sources consulted for drafting the template are included in the bibliography of this document [§ 6.]. They include, e.g.: ECP | Platform voor de InformatieSamenleving, 'Artificial Intelligence Impact Assessment', 2018, https://ecp.nl/wp-content/uploads/2019/01/Artificial-Intelligence-Impact-Assessment-English.pdf, (accessed on 27 February 2023); European Union Agency for Fundamental Rights, 'Getting the Future Right – Artificial Intelligence and Fundamental Rights', 2020, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf, (accessed on 27 February 2023); Madaio, M., et al., 'Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI', in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, 2020, paper 318; National Institute of Standards and Technology, 'Artificial Intelligence Risk Management Framework (AI RMF 1.0)', 2023, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf, (accessed on 27 February 2023); Numeum, 'Behind the Codes and the Data: A Practical Guide to Ethical AI', 2022, https://ai-ethical.com/wp-content/uploads/2022/12/2022-SN-Guide-Methodo-IA-Ethiques-web-version.pdf, (accessed on 27 February 2023).

[61] 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, AI Act Proposal, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, (accessed on 27 February 2023).

Regarding the proposed AI Act, some clarifications are here required. At the time of writing, the text of the Regulation has not been finalised and adopted. In light of the European Commission's proposal, it is reasonable to foresee that the Regulation will have an important impact on LEAs' technical capabilities, by prohibiting or restricting the use of some AI technologies creating unacceptable[62] or high risks.[63] However, for what concerns ethical principles and norms, the AI Act proposal adds little to what was already foreseen in the High-Level Expert Group's Ethics Guidelines for Trustworthy AI, especially when the relevant perspective is that of the end-user (and not that of the developer). Consequently, the AI Act proposal has limited relevance in updating the ethical principles already contained in MAGNETO's Ethical Risk Assessment Form.

**SIMPLIFY** – While the template needs to be detailed enough to allow LEAs to adequately address all the ethical concerns, it also needs to be understandable and not too complex. Considering the difficulties in predicting the likelihood of the ethical concerns to occur, the risk assessment approach needs to be abandoned in favour of an impact assessment one, which takes into account the severity of the prejudice suffered and the number of affected individuals.

The results of the revision of MAGNETO's Ethical Risk Assessment Form are incorporated in the ALIGNER AI System Governance template [§ 3.3.2].

## 3.2 Shaping a new fundamental rights impact assessment

A fundamental rights impact assessment (FRIA) can be defined as the "*process for identifying, understanding, assessing and addressing the adverse effects of a* […] *project or* […] *activities on the human rights enjoyment of impacted rights-holders*".[64]

The importance of performing a FRIA for AI systems has been widely underlined by international and national political institutions, scholars and civil society organisations. For instance, in its Assessment List for Trustworthy Artificial Intelligence, the High-Level Expert Group on AI underlined how a FRIA of AI systems should always be performed, even before conducting their ethical assessment.[65] More recently, the European Parliament proposed to amend the AI Act proposal to include an obligation to perform a FRIA.[66]

The most extensive guidance on how to perform a FRIA is given by the Danish Institute for Human Rights:[67] the Institute proposed a detailed 5-step methodology for evaluating the impact of business

---

[62] AI Act Proposal, Article 5, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, (accessed on 27 February 2023).
[63] AI Act Proposal, Article 6, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, (accessed on 27 February 2023).
[64] The Danish Institute for Human Rights, 'Welcome and introduction - Human rights impact assessment guidance and toolbox', 2020, https://www.humanrights.dk/sites/humanrights.dk/files/media/document/DIHR%20HRIA%20Toolbox_Welcome_and_Introduction_ENG_2020.pdf, (accessed on 27 February 2023), pp. 7-8.
[65] High-Level Expert Group on Artificial Intelligence, 'The Assessment List for Trustworthy Artificial Intelligence', 2020, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342, (accessed on 27 February 2023), p. 3.
[66] Euractiv, 'AI Act: MEPs want fundamental rights assessments, obligations for high-risk users', 2023, https://www.euractiv.com/section/artificial-intelligence/news/ai-act-meps-want-fundamental-rights-assessments-obligations-for-high-risk-users/, (accessed on 27 February 2023).
[67] The Danish Institute for Human Rights, 'Human rights impact assessment guidance and toolbox', 2020, https://www.humanrights.dk/tools/human-rights-impact-assessment-guidance-toolbox, (accessed on 27 February 2023).

projects on fundamental rights. AI-specific methodologies were also drafted by H. L. Janssen[68] and A. Mantelero.[69] These methodologies constitute the basis of a new operational tool, i.e., the ALIGNER Fundamental Rights Impact Assessment template [§ 3.3.1].

The ALIGNER Fundamental Rights Impact Assessment template aims to fully address the legal concerns related to the use of AI systems in a law enforcement domain. It adopts as benchmarks those fundamental rights susceptible to be infringed by LEAs; these are: the presumption of innocence and the right to an effective remedy and to a fair trial; the right to equality and non-discrimination; the freedom of expression and information; and the right to respect for private and family life and the right to protection of personal data.

The ALIGNER Fundamental Rights Impact Assessment template is not designed to be used by LEAs to fulfil their legal obligation to perform a DPIA. It is instead conceived as a complementary tool to a DPIA, expanding its scope to also assess and address the concerns related to the rights and freedoms of natural persons, as required by the LED.

## 3.3 Methodology of the ALIGNER Fundamental Rights Impact Assessment

The **ALIGNER Fundamental Rights Impact Assessment** (AFRIA) is a tool addressed to LEAs who aim to deploy AI systems for the purposes of prevention, investigation, detection or prosecution of criminal offences or execution of criminal penalties (i.e., law enforcement purposes) within the EU. As such, the AFRIA is **not** designed to be used in the following circumstances:

a. During the development stage of the AI systems, even if carried out by LEAs; and
b. When deploying AI systems for purposes other than law enforcement ones.

The AFRIA is a reflective exercise, seeking to further enhance the already existing legal and ethical governance systems of LEAs. Hence, the AFRIA has two main functions. First, it helps LEAs identify and mitigate the impact of the deployment of a certain AI system on ethical principles and (selected) fundamental rights of individuals. Second, it is a suitable instrument for LEAs to explain and record their decision-making processes. In other words, the AFRIA is a **process aimed to assist LEAs in building and demonstrating compliance with ethical principles and fundamental rights** while deploying AI systems in a law enforcement context.

### a. What the AFRIA addresses: A single AI system deployed for a single law enforcement purpose or a set of connected law enforcement purposes in a pre-determined context of use

An AFRIA addresses a **single AI system** deployed by LEAs. As a consequence, LEAs-users need to perform a separate AFRIA for each AI system they intend to deploy.

LEAs can perform a single AFRIA for an AI system deployed for either **a single law enforcement purpose or a set of connected law enforcement purposes**.[70] The connection between the purposes

[68] H. L. Janssen, 'An approach for a fundamental rights impact assessment to automated decision-making', *International Data Privacy Law*, vol. 10, no. 1, 2020, pp. 76-106.
[69] A. Mantelero, *Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI*, T.M.C. Asser Press The Hague, 2022.
[70] For instance, LEAs can perform a single AFRIA for an AI system deployed for both detection and prosecution of criminal offences.

needs to be evaluated in the particular case by the LEAs-users themselves. Therefore, it is of paramount importance for LEAs to always perform the AFRIA in relation to a **pre-determined context of use**. This may include, for instance, information on the AI system's target group, geographical area and time period of deployment, and trigger conditions.

#### b. When a AFRIA should be performed: Prior to the deployment of the AI system

In the EU, there is no legal obligation for LEAs deploying AI systems to perform an AFRIA, or an ethics and fundamental rights impact assessment in general. However, considering the particular sensitivity of the law enforcement domain, it is strongly advised to do so. As already seen above, an AFRIA complements the already existing legal and ethical governance systems of LEAs, as an instrument to further build and demonstrate the mandatory compliance with fundamental rights.

An AFRIA should be performed by LEAs **prior to the deployment of the AI system**, to inform the decision-making process on the *if*, *when*, *why* and *how* of the deployment. In case an AI system is already deployed for law enforcement purposes, LEAs are even more encouraged to conduct an AFRIA, unless their ethics and fundamental rights compliance was already and is currently evaluated via a similar instrument.

Performing an AFRIA is an iterative process. The AFRIA needs to be recorded, reviewed, and updated throughout the whole lifecycle of the AI system to reflect eventual changes in the functioning of the technology and/or its circumstances of deployment.

#### c. Who is responsible to perform the AFRIA: A dedicated multidisciplinary team

LEAs should establish a diverse and **multidisciplinary team**, responsible for performing the AFRIA. The team should include members of the organisation with legal, operational, and technical expertise. It is also advisable to involve the organisation's data protection officer in the AFRIA process.

If possible, LEAs should engage in discussions with the producer of the AI system assessed to clarify eventual uncertainties on the functioning of the AI system itself.

The AFRIA consists of two different, but connected, templates: the Fundamental Rights Impact Assessment [§ 3.3.1] and the AI System Governance [§ 3.3.2].

### 3.3.1 The Fundamental Rights Impact Assessment

The **Fundamental Rights Impact Assessment template** helps LEAs identify and assess the impact that the AI system they wish to deploy may have on the fundamental rights of individuals.

In ALIGNER D4.1, **four categories of fundamental rights** were identified as the most likely to be impacted by the use of AI systems in the law enforcement domain. These are:

1. Presumption of innocence and right to an effective remedy and to a fair trial;
2. Right to equality and non-discrimination;
3. Freedom of expression and information; and
4. Right to respect for private and family life and right to protection of personal data.

Accordingly, the Fundamental Rights Impact Assessment template is divided in four parts and, in each one of them, a group of fundamental rights is used as **benchmark for the following assessment**. To simplify the assessment process, the template contains an overview of the content of the four selected groups of fundamental rights, as defined by the CFREU [Figure 1].

**1. Presumption of innocence and right to an effective remedy and to a fair trial**

Everyone charged with a criminal offence must be presumed innocent until proved guilty according to law.
Everyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal.
Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights:

- to be informed promptly of the nature and cause of the accusation;
- to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and
- to obtain an adequately reasoned and accessible decision.

| Challenge | Evaluation | Estimated impact level |
|---|---|---|
| 1.1 The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision | | |
| 1.2 The AI system does not provide percentages or other indication on the degree of likelihood that the outcome is correct/incorrect, prejudicing the user that there is no possibility of error and therefore that the outcome is undoubtedly incriminating | | |
| 1.3 The AI system produces an outcome that forces a reversal of burden of proof upon the suspect, by presenting itself as an absolute truth, practically depriving the defence of any chance to counter it | | |
| 1.4 There is no explanation of reasons and criteria behind a certain output of the AI system that the user can understand | | |
| 1.5 There is no indication of the extent to which the AI system influences the overall decision-making process | | |

Figure 1: Example of Fundamental Rights Impact Assessment template, emphasis added

### a. ‘Challenge’ column

To help and guide LEAs-users in their assessment, the template already lists some ‘**challenges**’. These are some possible **characteristics embedded in AI systems that may have a negative impact on the fundamental right** [Figure 2]. The challenges are formulated in a negative form (e.g., “*there is no* …”), so as to reduce the risk of acquiescence biases and stimulate further thought. LEAs may rely on the pre-listed challenges or add additional ones, as required.

| 1. Presumption of innocence and right to an effective remedy and to a fair trial | | |
|---|---|---|
| Everyone charged with a criminal offence must be presumed innocent until proved guilty according to law.<br>Everyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal.<br>Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights:<br>❖ to be informed promptly of the nature and cause of the accusation;<br>❖ to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and<br>to obtain an adequately reasoned and accessible decision. | | |
| Challenge | Evaluation | Estimated impact level |
| 1.1 The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision | | |
| 1.2 The AI system does not provide percentages or other indication on the degree of likelihood that the outcome is correct/incorrect, prejudicing the user that there is no possibility of error and therefore that the outcome is undoubtedly incriminating | | |
| 1.3 The AI system produces an outcome that forces a reversal of burden of proof upon the suspect, by presenting itself as an absolute truth, practically depriving the defence of any chance to counter it | | |
| 1.4 There is no explanation of reasons and criteria behind a certain output of the AI system that the user can understand | | |
| 1.5 There is no indication of the extent to which the AI system influences the overall decision-making process | | |

Figure 2: Example of Fundamental Rights Impact Assessment template, emphasis added

### b. 'Evaluation' column

In the '**evaluation**' column, LEAs need to identify **how the listed challenges relate to the assessed AI system**, for the identified law enforcement purposes and in relation to the envisaged context of use. In other words, LEAs need to explain both *whether* and, if so, *to what degree,* the assessed AI system embeds each of the challenges, and *how* it does so [Figure 3].

| 1. Presumption of innocence and right to an effective remedy and to a fair trial | | |
|---|---|---|
| Everyone charged with a criminal offence must be presumed innocent until proved guilty according to law.<br>Everyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal.<br>Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights:<br>❖ to be informed promptly of the nature and cause of the accusation;<br>❖ to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and<br>to obtain an adequately reasoned and accessible decision. | | |
| Challenge | Evaluation | Estimated impact level |
| 1.1 The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision | The AI system communicates that the outcome is the result of an algorithmic decision only in case of flagging of an individual, while the communication is omitted in case of no flag | |
| 1.2 The AI system does not provide percentages or other indication on the degree of likelihood that the outcome is correct/incorrect, prejudicing the user that there is no possibility of error and therefore that the outcome is undoubtedly incriminating | The AI system does not communicate the likelihood of the output and it is impossible for the user to establish it | |
| 1.3 The AI system produces an outcome that forces a reversal of burden of proof upon the suspect, by presenting itself as an absolute truth, practically depriving the defence of any chance to counter it | When the AI system flags an individual, a further investigation against them is immediately started, even in absence of other evidence incriminating the subject | |
| 1.4 There is no explanation of reasons and criteria behind a certain output of the AI system that the user can understand | The AI system does not communicate the user the reasons and criteria behind any of the output reached and the user cannot understand them with any other means | |
| 1.5 There is no indication of the extent to which the AI system influences the overall decision-making process | The weight of the output of the AI system in the overall decision-making process was not specifically evaluated | |

Figure 3: Example of Fundamental Rights Impact Assessment template, emphasis and text added

### c. ‘Estimated impact’ column

In the ‘**estimated impact**’ column, LEAs need to estimate the level of the **negative effect** the deployment of the AI system may have on the fundamental right of individuals, due to the already evaluated challenges posed by the AI system’s characteristics. In doing so, LEAs need to consider the following factors:

1. the severity of prejudice, namely how serious is the prejudice experienced by the affected individuals; and
2. the number of affected individuals.

The impact matrix below helps the user estimate and visualize impacts.

| | | **Severity of prejudice** | | |
|---|---|---|---|---|
| | | **Negligible**<br>Affected individuals may experience no prejudice | **Critical**<br>Affected individuals may experience prejudice | **Catastrophic**<br>Affected individuals may experience a serious prejudice |
| **Number of affected individuals** | **Low**<br>The percentage of people affected is small | Low | Low | Medium |
| | **Medium**<br>Whilst the absolute number of people affected is small, a vulnerable group is particularly impacted | Low | Medium | High |
| | **High**<br>The percentage of people affected is significant | Medium | High | Very high |

Table 1: Impact matrix

The user should estimate both the severity of the prejudice (in *negligible*, *critical,* or *catastrophic*) and the number of affected individuals (in *low*, *medium,* or *high*). Based on the estimations, the user finds the impact level (*low*, *medium*, *high,* or *very high*) in the square where the severity of the prejudice and the number of affected individuals meet.

For instance, in relation to challenge 1.1, if the user estimates the severity of the prejudice as *critical* and the number of affected individuals as *medium*, the impact level will be *medium* [Figure 4].

| 1. Presumption of innocence and right to an effective remedy and to a fair trial | | |
|---|---|---|
| Everyone charged with a criminal offence must be presumed innocent until proved guilty according to law.<br>Everyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal.<br>Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights:<br>❖ to be informed promptly of the nature and cause of the accusation;<br>❖ to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and<br>to obtain an adequately reasoned and accessible decision. | | |
| Challenge | Evaluation | Estimated impact level |
| 1.1 The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision | The AI system communicates that the outcome is the result of an algorithmic decision only in case of flagging of an individual, while the communication is omitted in case of no flag | Medium |
| 1.2 The AI system does not provide percentages or other indication on the degree of likelihood that the outcome is correct/incorrect, prejudicing the user that there is no possibility of error and therefore that the outcome is undoubtedly incriminating | The AI system does not communicate the likelihood of the output and it is impossible for the user to establish it | High |
| 1.3 The AI system produces an outcome that forces a reversal of burden of proof upon the suspect, by presenting itself as an absolute truth, practically depriving the defence of any chance to counter it | When the AI system flags an individual, a further investigation against them is immediately started, even in absence of other evidence incriminating the subject | Very high |
| 1.4 There is no explanation of reasons and criteria behind a certain output of the AI system that the user can understand | The AI system does not communicate the user the reasons and criteria behind any of the output reached and the user cannot understand them with any other means | Very high |
| 1.5 There is no indication of the extent to which the AI system influences the overall decision-making process | The weight of the output of the AI system in the overall decision-making process was not specifically evaluated | High |

Figure 4: Example of Fundamental Rights Impact Assessment template, emphasis and text added

### 3.3.2 The AI System Governance

The **AI System Governance** template helps LEAs identify, explain, and record possible measures to mitigate the negative impact that the deployment of the AI system would have on the ethical principles and the fundamental rights of individuals.

In 2019, The High-Level Expert Group on Artificial Intelligence set up by the European Commission published its 'Ethics Guidelines for Trustworthy AI'.[71] There, the Group identified seven key requirements that an AI system should fulfil to be considered 'trustworthy', i.e., a lawful, ethical, and robust AI system. These requirements are:

1. Human agency and oversight;
2. Technical robustness and safety;
3. Privacy and data governance;
4. Transparency;
5. Diversity, non-discrimination and fairness;
6. Societal and environmental wellbeing; and
7. Accountability.

Accordingly, the AI system Governance template is divided in seven parts and, in each one of them, a key requirement for trustworthy AI is used as **benchmark for grouping the minimum standards** that an AI system should achieve [Figure 5].

[71] High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, (accessed on 8 February 2023).

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | | | | | | |
| | | [1.5] | | | | | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | | | | | | |
| | | [2.2] | | | | | | |
| | | [4.1] | | | | | | |
| Human oversight | ❒ The AI system does not affect human autonomy by interfering with the user decision-making process | [1.2] | | | | | | |
| | | [1.3] | | | | | | |
| | | [1.5] | | | | | | |
| | | [4.1] | | | | | | |

Figure 4: Example of AI System Governance template, emphasis added

### a. ‘Component’ column

In the ‘**component**’ column, the **building blocks substantiating the considered key requirement** are listed [Figure 6].

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | | | | | | |
| | | [1.5] | | | | | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | | | | | | |
| | | [2.2] | | | | | | |
| | | [4.1] | | | | | | |
| Human oversight | ❒ The AI system does not affect human autonomy by interfering with the user decision-making process | [1.2] | | | | | | |
| | | [1.3] | | | | | | |
| | | [1.5] | | | | | | |
| | | [4.1] | | | | | | |

Figure 5: Example of AI System Governance template, emphasis added

### b. ‘Minimum standards to be achieved’ column

To help and guide LEAs-users in their decision-making process, the template already lists some ‘**minimum standards to be achieved**’. These are some possible **characteristics that an AI system should embed** or possible **governance procedures that the organisation should always implement for the deployment of the AI system to be considered trustworthy** [Figure 7].

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | | | | | | |
| | | [1.5] | | | | | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | | | | | | |
| | | [2.2] | | | | | | |
| | | [4.1] | | | | | | |

Figure 6: Example of AI System Governance template, emphasis added

### c. ‘Initial impact estimate’ column

To further help and guide LEAs-users in their decision-making process, in the ‘**initial impact estimate**’ column, the template already connects the **minimum standard with** (at least) **one previously estimated challenge and impact level**, as that was already estimated in the Fundamental Rights Impact Assessment template. The link between the minimum standard and the estimated impact is highlighted where the minimum standards are suitable to mitigate possible negative impacts that the deployment of the AI system would have on the fundamental rights of the individuals. The numbers (e.g., 1.2, 1.5, and so on) correspond to the ‘challenges’ listed in the Fundamental Rights Impact Assessment template. For each of the challenges, LEAs need to manually report the impact level (i.e., *low*, *medium*, *high,* or *very high*), as it was already estimated in the Fundamental Rights Impact Assessment template [Figure 8].

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | High | | | | | |
| | | [1.5] | High | | | | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | High | | | | | |
| | | [2.2] | Low | | | | | |
| | | [4.1] | Very high | | | | | |

Figure 7: Example of AI System Governance template, emphasis and text added

Where the minimum standards are not suitable to mitigate possible negative impacts that the deployment of the AI system would have on the fundamental rights of the individuals, the '**initial impact estimate**' column is **left blank** [Figure 9].

| 2. Transparency | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Traceability | ❒ There are mechanisms to ensure the traceability of the input data used by the AI system and its outcomes | | | | | | | |

Figure 8: Example of AI System Governance template, emphasis added

### d. 'Additional mitigation measures implemented' column

Whenever an initial impact is linked to a minimum standard, in the '**additional mitigation measures implemented**' column, LEAs need to state:

- ***if* and *how*** the minimum standard **is (foreseen to be) implemented** in the AI system and/or within the organisation; and
- ***how*** the minimum standard **is suitable to mitigate the connected previously estimated impact**, by paying particular attention to how the standard is reducing the severity of the prejudice and/or the number of affected individuals [Figure 10].

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | High | The AI system reveals the likelihood of the output, so that the user can take an informed decision on the follow-up actions | | | | |
| | | [1.5] | High | The user can play an active role in the decision-making process, by modifying the parameters informing the decision of the AI system | | | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | High | The weight of the output of the AI system in the decision-making processes of the organisation is concretely evaluated. The results are made known to the users, who are tasked to take an informed decision on the follow-up actions | | | | |

Figure 9: Example of AI System Governance template, emphasis and text added

Whenever an initial impact is not linked to a minimum standard, and thereby left blank, in the '**additional mitigation measures implemented**' column, LEAs need to state:

- ***if* and *how*** the minimum standard **is (foreseen to be) implemented** in the AI system and/or within the organisation [Figure 11].

| 2. Transparency | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Traceability | ❒ There are mechanisms to ensure the traceability of the input data used by the AI system and its outcomes | | | Both the input data and the outcomes are recorded and accessible to the user | | | | |

Figure 10: Example of AI System Governance template, emphasis and text added

### e. 'Final assessment' column

Whenever an initial impact is **linked** to a minimum standard, in the '**final assessment**' column, LEAs need to:

- Use the impact matrix seen above [Table 1], to estimate the **final impact level** on fundamental rights that the deployment of the AI system may have, despite the implementation of additional mitigation measures; and
- if any, list **further actions** suitable to improve the implementation of the minimum standard and further mitigate the final impact on fundamental rights, for instance in case where the mitigation measures are not considered sufficient in relation to the estimated impact [Figure 12].

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | High | The AI system reveals the likelihood of the output, so that the user can take an informed decision on the follow-up actions | Low | | | |
| | | [1.5] | High | The user can play an active role in the decision-making process, by modifying the parameters informing the decision of the AI system | Medium | Implementing a mechanism to allow the user to add new parameters informing the decision of the AI system | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | High | The weight of the output of the AI system in the decision-making processes of the organisation is concretely evaluated. The results are made known to the users, who are tasked to take an informed decision on the follow-up actions | Low | | | |

Figure 11: Example of AI System Governance template, emphasis and text added

Whenever an initial impact is **not linked** to a minimum standard, in the '**final assessment**' column, LEAs need to:

- list, if any, **further actions** suitable to improve the implementation of the minimum standard and further mitigate the final impact on fundamental rights, for instance in case where the mitigation measures are not considered sufficient in relation to the estimated impact [Figure 13].

| 2. Transparency | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Traceability | ❒ There are mechanisms to ensure the traceability of the input data used by the AI system and its outcomes | | | Both the input data and the outcomes are recorded and accessible to the user | | | | |

Figure 12: Example of AI System Governance template, emphasis and text added

### f. 'Responsible department' and 'timeline' columns

In the '**responsible department**' and '**timeline**' column, LEAs need to specify the department of their organisation responsible for the implementation of the mitigation measures foreseen, and their (estimated) timeline of adoption [Figure 14].

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | High | The AI system reveals the likelihood of the output, so that the user can take an informed decision on the follow-up actions | Low | | ICT department | June 2023 |
| | | [1.5] | High | The user can play an active role in the decision-making process, by modifying the parameters informing the decision of the AI system | Medium | Implementing a mechanism to allow the user to add new parameters informing the decision of the AI system | ICT and legal departments | March 2023 |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | High | The weight of the output of the AI system in the decision-making processes of the organisation is concretely evaluated. The results are made known to the users, who are tasked to take an informed decision on the follow-up actions | Low | | ICT and legal departments | March 2023 |

Figure 13: Example of AI System Governance template, emphasis and text added

# 4. ALIGNER Fundamental Rights Impact Assessment template

The ALIGNER Fundamental Rights Impact Assessment template assists LEAs in building and demonstrating compliance with ethical principles and fundamental rights while deploying AI systems in law enforcement context.

## 4.1 How to use the AFRIA Template

1. LEAs-users should use the AFRIA template as part of their legal and ethical governance systems. It is important to specify the '**AI system assessed**' and give a '**detailed description of the technology and input data**' and a '**detailed description of the purposes and context of use**'.
2. In Template #1 (Fundamental Right Impact Assessment), **four groups of fundamental rights** are used as benchmark for the following assessment.
3. In Template #1, the user finds a list of '**challenges**'. These are some possible characteristics embedded in the AI systems that may have a negative impact on the fundamental right.
4. In Template #1, the user should precise in the '**evaluation'** column how the challenges relate to the assessed AI system, so *whether*, and, if so, *to what degree*, the assessed AI system embeds the considered challenge, and *how* it does so.
5. In Template #1, using the impact matrix below, the user should estimate the level of the '**estimated impact**' that the deployment of the AI system would have on the fundamental right.
6. In Template #2 (AI System Governance), the **7 key requirements** for trustworthy AI identified from the High-Level Expert Group on AI are used as benchmark for grouping the minimum standards that an AI system should achieve.
7. In Template #2, the user finds a list of '**minimum standards to be achieved**'. If suitable to mitigate possible negative impacts that the deployment of the AI system would have on the fundamental rights of individuals, the minimum standards are linked to one previously estimated impact in the '**initial impact estimate**' column– otherwise the latter column is left blank.
8. In Template #2, the user should precise in the '**additional mitigation measures implemented**' column *if* and *how* the minimum standard is (foreseen to be) implemented and *how* it is suitable to mitigate the connected estimated impact.
9. In Template #2, using the impact matrix below, the user should estimate the level of the '**final estimated impact**' on fundamental rights and the '**further actions**' needed to achieve the minimum standards and further mitigate the final estimated impact.

| | | Severity of prejudice | | |
|---|---|---|---|---|
| | | **Negligible**<br>Affected individuals may experience no prejudice | **Critical**<br>Affected individuals may experience prejudice | **Catastrophic**<br>Affected individuals may experience a serious prejudice |
| **Number of affected individuals** | **Low**<br>The percentage of people affected is small | Low | Low | Medium |
| | **Medium**<br>Whilst the absolute number of people affected is small, a vulnerable group is particularly impacted | Low | Medium | High |
| | **High**<br>The percentage of people affected is significant | Medium | High | Very high |

Table 2: Impact matrix

## 4.2 Template #1: Fundamental Rights Impact Assessment

| Fundamental Rights Impact Assessment Template | |
|---|---|
| Name | |
| Organisation/Position | |
| Date | |
| Contributors | |
| AI system assessed | |
| Detailed description of the technology and input data | |
| Detailed description of the purposes and context of use | |

| 1. Presumption of innocence and right to an effective remedy and to a fair trial | | |
|---|---|---|
| Everyone charged with a criminal offence must be presumed innocent until proved guilty according to law.<br>Everyone whose rights and freedoms are violated has the right to an effective remedy before a tribunal.<br>Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law, including rights:<br>❖ to be informed promptly of the nature and cause of the accusation;<br>❖ to bring their arguments and evidence as well as scrutinise and counteract the evidence presented against them; and<br>to obtain an adequately reasoned and accessible decision. | | |
| Challenge | Evaluation | Estimated impact level |
| 1.1 The AI system does not communicate that a decision/advice or outcome is the result of an algorithmic decision | | |
| 1.2 The AI system does not provide percentages or other indication on the degree of likelihood that the outcome is correct/incorrect, prejudicing the user that there is no possibility of error and therefore that the outcome is undoubtedly incriminating | | |
| 1.3 The AI system produces an outcome that forces a reversal of burden of proof upon the suspect, by presenting itself as an absolute truth, practically depriving the defence of any chance to counter it | | |
| 1.4 There is no explanation of reasons and criteria behind a certain output of the AI system that the user can understand | | |
| 1.5 There is no indication of the extent to which the AI system influences the overall decision-making process | | |
| 1.6 There is no set of measures that allow for redress in case of the occurrence of any harm or adverse impact | | |

| 2. Right to equality and non-discrimination | | |
|---|---|---|
| Everyone is equal before the law.<br>Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited.<br>Everyone should be protected against discriminatory decisions or policies, including automated decision-making based on sensitive data. | | |
| Challenge | Evaluation | Estimated impact level |
| 2.1 The AI system targets members of a specific social group | | |
| 2.2 There are no mechanisms to flag and correct issues related to bias, discrimination, or poor performance | | |
| 2.3 The AI system does not consider the diversity and representativeness for specific population or problematic use cases | | |
| **3. Freedom of expression and information** | | |
| Everyone has the right to freedom of expression, including freedom to hold opinions, communicate and acquire information<br>❖ State negative obligation not to interfere and positive obligation to facilitate the exercise of the right | | |
| Challenge | Evaluation | Estimated impact level |
| 3.1 There is no mechanism to limit the deployment of the AI system to suspected individuals | | |
| 3.2 The data stored, recorded, and produced are not easily accessible to concerned individuals | | |

| 4. Right to respect for private and family life and right to protection of personal data | | |
|---|---|---|
| Everyone has the right to respect for their private and family life, home and communications.<br>❖ Self-development without state interference.<br>❖ Everyone has the right to the protection of personal data concerning them.<br>❖ Personal data must be processed fairly for specified purposes and on a legitimate basis.<br>❖ Rights of access and rectification.<br>❖ Independent oversight. | | |
| Challenge | Evaluation | Estimated impact level |
| 4.1 There are no mechanisms for the user to exercise control over the processing of personal data | | |
| 4.2 There are no measures to ensure the lawfulness of the processing of personal data | | |
| 4.3 There are no procedures to limit the access to personal data and to the extent and amount necessary for those purposes | | |
| 4.4 There is no mechanism allowing to comply with the exercise of data subject's rights (access, rectification and erasure of data relating to a specific individual) | | |
| 4.5 There are no specific measures in place to enhance the security of the processing of personal data (via encryption, anonymisation and aggregation) | | |
| 4.6 There is no procedure to conduct a data protection impact assessment | | |

## 4.3 Template #2: AI System Governance

| AI System Governance Template | |
|---|---|
| Name | |
| Organisation /Position | |
| Date | |
| Contributors | |
| AI system | |
| Detailed description of the technology and input data | |
| Detailed description of the purposes and context of use | |

| 1. Human autonomy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Human agency | ❒ The task allocation between the AI system and the user allows meaningful interactions | [1.2] | | | | | | |
| | | [1.5] | | | | | | |
| | ❒ There are procedures to describe the level of human involvement and the moments for human interventions | [1.5] | | | | | | |
| | | [2.2] | | | | | | |
| | | [4.1] | | | | | | |
| Human oversight | ❒ The AI system does not affect human autonomy by interfering with the user decision-making process | [1.2] | | | | | | |
| | | [1.3] | | | | | | |
| | | [1.5] | | | | | | |
| | | [4.1] | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Human oversight | ❒ There are mechanisms to prevent overconfidence or over-reliance in the results offered by the AI system | [1.1] | | | | | | |
| | | [1.2] | | | | | | |
| | ❒ There are mechanisms to detect and correct wrong outputs | [1.6] | | | | | | |
| | | [2.2] | | | | | | |
| | | [2.3] | | | | | | |
| | ❒ There are mechanisms to safely abort an entire operation when needed | | | | | | | |

| **2. Transparency** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Traceability | ❒ There are mechanisms to ensure the traceability of the input data used by the AI system and its outcomes | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Explainability | ❒ It is possible for the user to understand and explain the reasons and criteria behind a certain output of the AI system | [1.4] | | | | | | |
| Communication | ❒ There are procedures enabling the user to communicate to the public that decisions are taken on the basis of an algorithmic process | [1.3] | | | | | | |
| | ❒ There are procedures enabling the user to explain to the public the purposes, characteristics, limitations, and shortcomings of the AI system | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Communication | ❒ There are procedures enabling the user to make the data stored, recorded, and produced available to concerned individuals | [3.2] | | | | | | |
| **3. Diversity, non-discrimination and fairness** | | | | | | | | |
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Unfair bias avoidance | ❒ There are procedures to test and evaluate the diversity and representativeness of the used datasets, also for specific social group or use cases | [2.3] | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Unfair bias avoidance | ❒ There are procedures to test and evaluate the diversity and representativeness of the algorithm used, also for specific social groups or use cases | [2.3] | | | | | | |
| | ❒ There are procedures to evaluate whether specific social groups are disproportionally affected by the AI system | [2.1] | | | | | | |
| | ❒ There are mechanisms to flag and correct bias, discrimination or poor performance | [2.2] | | | | | | |

| 4. Democracy and societal wellbeing | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Social impact | ❒ There are procedures to ensure that the social impacts of the AI systems are well understood by the public | | | | | | | |
| Society and democracy | ❒ There are procedures to assess the broad social impact of the AI system (e.g., chilling effect, power asymmetry, trust, …) | | | | | | | |
| | ❒ There are mechanisms to limit the deployment of the AI system to groups of individuals on the basis of suspicion/objective criteria | [3.1] | | | | | | |

| 5. Privacy and data governance | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact Level | | Final estimated impact level | Further actions | | |
| Respect for privacy and data protection | ❒ There are mechanisms for the user to exercise control over the processing of personal data | [4.1] | | | | | | |
| | ❒ There are measures to ensure the lawfulness of the processing of personal data | [4.2] | | | | | | |
| | ❒ There are measures to minimise the amount of personal data processed | [4.3] | | | | | | |
| | ❒ There is a mechanism allowing to comply with data subjects' rights | [4.4] | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Quality and integrity of data | ❒ There are specific measures to enhance the security of the processing of personal data (via encryption, anonymization and aggregation) | [4.5] | | | | | | |
| | ❒ There are processes to ensure the quality and integrity of data | | | | | | | |
| | ❒ The AI system is aligned with relevant standards (ISO, IEEE) for data security, management and governance | | | | | | | |
| Access to data | ❒ There are procedures to limit the access to personal data | [4.3] | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Governance | ❒ There is a procedure to conduct a data protection impact assessment | [4.6] | | | | | | |
| | ❒ A data protection officer has been appointed | | | | | | | |
| | ❒ There are mechanisms to allow reporting of processing activities to the supervisory body | | | | | | | |
| International data transfers | ❒ There are mechanisms to control the transfer of personal data to third countries | | | | | | | |

| **6. Technical robustness and safety** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Security | ❒ The potential vulnerability of the AI system has been assessed | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Security | ❒ There are mechanisms to ensure the integrity and resilience of the AI system against potential cyberattacks | | | | | | | |
| Fallback and general safety | ❒ There is a fallback plan for adversarial attacks or unexpected situations | | | | | | | |
| Accuracy | ❒ There is an assessment of the level of accuracy required in relation to the envisaged use | | | | | | | |
| | ❒ There are mechanisms to evaluate and ensure that the used datasets are comprehensive and up to date | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Reliability and reproducibility | ❒ There are procedures to evaluate the reliability and reproducibility of the AI system's aspects (inputs and outputs), also in specific contexts | | | | | | | |
| **7. Accountability** | | | | | | | | |
| Component | Minimum standards to be achieved | Initial impact estimate | | Additional mitigation measures implemented | Final assessment | | Responsible department | Timeline |
| | | Challenge no. | Impact level | | Final estimated impact level | Further actions | | |
| Competence | ❒ There are clear programs to provide information on the role of the operator, the competencies required to operate the AI system and the implications of operator error | | | | | | | |
| | ❒ There are safeguards against incompetent operation of the AI system | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Misuse awareness | ❒ There is an assessment of the likelihood of misuse of the AI system and of its possible outcomes | | | | | | | |
| | ❒ There are ethics education and security awareness programs to sensitise the users to the potential risk of misuse | | | | | | | |
| Auditability | ❒ There are legged and traceable procedures to enable independent audit, also in order to remedy to identified issues in the AI system | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Ability to redress | ❒ There are measures that allow redress in case of the occurrence of any harm or adverse impact | [1.6] | | | | | | |
| | ❒ There are procedures to provide information to affected parties about opportunity for redress | | | | | | | |

# 5. Validation of the ALIGNER FRIA

To validate the ALIGNER FRIA two separate interactive workshop sessions were organised.

The first validation session, titled 'Ideas for ethical and legal impact assessment', was part of ALIGNER Workshop no. 4 (30th November – 1st December 2022, online) and gathered feedback on the initial ideas for the ethical and impact assessment. The participants were the members of ALIGNER Advisory Boards, the Law Enforcement Agency Advisory Board (LEAAB) and the Scientific, Industrial and Ethical Advisory Board (SIEAB). During the session, the participants were separated in three groups and each group received a copy of one of the following templates: a DPIA, MAGNETO's Ethical Risk Assessment and MAGNETO's Misuse Risk Assessment. Participants were asked to fill-in the templates based on the first ALIGNER scenario (i.e., disinformation and social manipulation) and to identify gaps and possible improvements.

The second validation session, titled 'Ensuring compliance during use – The ALIGNER methodology' was part of the 'Ethical and Legal AI for Security' conference organised by the SU-AI H2020 projects ALIGNER, STARLIGHT and popAI (25th – 26th January 2023, Brussels). The participants were national Law Enforcement Agencies, researchers, civil society, ethicists, legal and social experts, industry, policy makers and European Agencies. During the session, the participants were presented with the intermediate ALIGNER FRIA templates and asked to provide feedback on the methodology and usability for LEAs.

The participants in both workshop sessions helped to identify the methodological and execution gaps, offered suggestions for improvement of the FRIA form, and supported the identification of specific solutions.

After the completion of the ALIGNER FRIA templates, the final draft was sent to the members of the ALIGNER Advisory Boards which have been asked to provide feedback. The FRIA templates were deemed useful by the participants who consider them "*complete and accurate*" and helpful to create and implement a legal and ethical governance system for AI systems used in the law enforcement domain. The participants also made suggestions to improve the form:

1. Create a shorter form with a complete example at the beginning of the document;
2. Add a list of relevant EU laws/directives/guidelines/best practices on the respective subjects;
3. Implement more ethical and legal training, including the use of FRIA.

The final version of the ALIGNER FRIA templates, integrating the feedbacks of the ALIGNER Advisory Boards, was presented during popAI's Plenary Meeting (14th March 2023, Rome) and was well received from the audience.

# 6. Bibliography

## Legislation

Charter of Fundamental Rights of the European Union, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012P%2FTXT, (accessed on 27 February 2023).

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, and repealing Council Framework Decision 2008/977/JHA, LED, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680, (accessed on 20 February 2023).

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, AI Act Proposal, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, (accessed on 27 February 2023).

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), GDPR, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679, (accessed on 20 February 2023).

## Literature

Beauchamp, T. L., and Childress, J. F., *Principles of Biomedical Ethics*, Oxford University Press USA, 2013.

ECP | Platform voor de InformatieSamenleving, 'Artificial Intelligence Impact Assessment', 2018, https://ecp.nl/wp-content/uploads/2019/01/Artificial-Intelligence-Impact-Assessment-English.pdf, (accessed on 27 February 2023).

Eren, E., Casaburo, D., and Vogiatzoglou, P., 'ALIGNER D4.1 – State-of-the-art reports on ethics & law aspects in Law Enforcement and Artificial Intelligence', 2022.

European Union Agency for Fundamental Rights, 'Getting the Future Right – Artificial Intelligence and Fundamental Rights', 2020, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf, (accessed on 27 February 2023).

Guelke, H., 'SURVEILLE D2.2 – Paper with input from end-users', 2013, https://surveille.eui.eu/wp-content/uploads/sites/19/2015/04/D2.2.Paper-with-Input-from-End-Users.pdf, (accessed on 25 February 2023).

Harris, I., *et al.*, 'Ethical assessment of new technologies: a meta-methodology', *Journal of Information, Communication & Ethics in Society*, vol. 9, no. 1, 2011, pp. 49-64.

Hartz Sørake, J., and Brey, P., 'SATORI D1.1 – Annex 2.b.1: Ethics Assessment in Different Field – Information Technologies, 2015, https://satoriproject.eu/media/2.b.1-Information-technology.pdf, (accessed on 22 February 2023).

High-Level Expert Group on Artificial Intelligence, 'The Assessment List for Trustworthy Artificial Intelligence', 2020, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342, (accessed on 27 February 2023).

High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, (accessed on 27 February 2023).

Janssen, H. L., 'An approach for a fundamental rights impact assessment to automated decision-making', *International Data Privacy Law*, vol. 10, no. 1, 2020, pp. 76-106.

Kloza, D., *et al*., 'Data protection impact assessment in the European Union: developing a template for a report from the assessment process', *d.pia.lab Policy Brief No. 1/2020*, 2020, https://cris.vub.be/ws/portalfiles/portal/53602836/dpialab_pb2020_1_final.pdf, (accessed on 20 February 2023).

Madaio, M., *et al*., 'Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI', in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, 2020, paper 318.

Mantelero, A., *Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI*, T.M.C. Asser Press The Hague, 2022.

Marquenie, T., and Quezada-Tavárez, K., 'Data Protection Impact Assessments in Law Enforcement', in G. Markarian *et al*. (eds.), *Security Technologies and Social Implications*, Wiley-IEEE Press, 2023, pp. 32-60.

Marquenie, T., *et al*., 'MAGNETO D9.3 – Interim Ethical and Legal Assessment', 2019, https://results.magneto-h2020.eu/wp-content/uploads/2021/04/D9.3_Interim_Ethical_and_Legal_Assessment_compressed.pdf, (accessed on 24 February 2023)

Marsh, I., Hale, N., and Kelly, D., *Ethical Assessment Regarding the Use or Misuse of AI Systems for Law Enforcement. A Handbook for Law Enforcement Officials*, 2021.

National Institute of Standards and Technology, 'Artificial Intelligence Risk Management Framework (AI RMF 1.0)', 2023, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf, (accessed on 27 February 2023).

Numeum, 'Behind the Codes and the Data: A Practical Guide to Ethical AI', 2022, https://ai-ethical.com/wp-content/uploads/2022/12/2022-SN-Guide-Methodo-IA-Ethiques-web-version.pdf, (accessed on 27 February 2023).

Reijers, W., *et al*., 'SATORI D4.1 – Annex 1: A Common Framework for Ethical Impact Assessment', 2016, https://satoriproject.eu/media/D4.1_Annex_1_EIA_Proposal.pdf, (accessed on 22 February 2023).

Shelley-Egan, C., *et al.*, 'SATORI D1.1 – Ethical Assessment of Research and Innovation: A Comparative Analysis of Practices and Institutions in the EU and selected other countries', 2016, https://satoriproject.eu/media/D1.1_Ethical-assessment-of-RI_a-comparative-analysis-1.pdf, (accessed on 22 February 2023).

The Danish Institute for Human Rights, 'Welcome and introduction - Human rights impact assessment guidance and toolbox', 2020, https://www.humanrights.dk/sites/humanrights.dk/files/media/document/DIHR%20HRIA%20Toolbox_Welcome_and_Introduction_ENG_2020.pdf, (accessed on 27 February 2023).

Wright, D., 'Ethical Impact Assessment', in J. Britt Holbrook and C. Mitcham (eds.), *Ethics, Science, Technology and Engineering: A Global Resource*, Macmillan Reference USA, 2014.

Wright, D., and Mordini, E., 'Privacy and Ethical Impact Assessment', in D. Wright and P. De Hert, *Privacy Impact Assessment*, Springer, 2012, pp. 397-418.

Wright, D., 'A framework for the ethical impact assessment of information technology', *Ethics and Information Technology*, vol. 13, no. 3, 2011, pp.199-226.

## Soft-law

Agencia Española Protección Datos, 'Template for Data Protection Impact Assessment Report (DPIA) For Public Adminisitrations', 2022, https://www.aepd.es/es/documento/modelo-informe-EIPD-AAPP-en.rtf, (accessed on 20 February 2023).

Agencia Española Protección Datos, 'Template for Data Protection Impact Assessment Report (DPIA) For Private Sector', 2022, https://www.aepd.es/es/documento/modelo-informe-EIPD-sector-privado-en.rtf, (accessed on 20 February 2023).

Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', 2017, https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711, (accessed on 20 February 2023).

Commission Nationale de l'Informatique et des Libertés, 'Privacy Impact Assessment (PIA)', https://www.cnil.fr/en/privacy-impact-assessment-pia, (accessed on 20 February 2023).

Information and Data Protection Commissioner, 'Guidelines on DPIA template', https://idpc.org.mt/wp-content/uploads/2020/07/Guidelines-on-DPIA-template.pdf, (accessed on 20 February 2023).

Information Commissioner's Office, 'Sample DPIA template', June 2018, https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fico.org.uk%2Fmedia%2Ffor-organisations%2Fdocuments%2F2553993%2Fdpia-template.docx&wdOrigin=BROWSELINK, (accessed on 20 February 2023).

Information Commissioner's Office, 'Data Protection Impact Assessments (DPIAs)', https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/, (accessed on 20 February 2023).

Information Commissioner's Office, 'Guide to Law Enforcement Processing', https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-le-processing/, (accessed on 20 February 2023).

Information Commissioner's Office, 'Toolkit for organizations considering using data analytics', https://ico.org.uk/for-organisations/toolkit-for-organisations-considering-using-data-analytics/, (accessed on 20 February 2023).

Information Commissioner of the Republic of Slovenia, 'Privacy Impact Assessment (PIA) Guidelines for the Introduction of new Police Powers', 2014 https://www.ip-rs.si/fileadmin/user_upload/Pdf/smernice/PIA_guideliness_for_introduction_of_new_police_powers_english.pdf, (accessed on 20 February 2023).

## Other sources

ethicstoolkit.ai, 'Ethics & Algorithms Toolkit', https://ethicstoolkit.ai/, (accessed on 27 February 2023).

Euractiv, 'AI Act: MEPs want fundamental rights assessments, obligations for high-risk users', 2023, https://www.euractiv.com/section/artificial-intelligence/news/ai-act-meps-want-fundamental-rights-assessments-obligations-for-high-risk-users/, (accessed on 27 February 2023).

SATORI, 'Work Package 1:Comparative Analysis of Ethics Assessment Practices – Deliverable 1.1: Ethical Assessment of R&I: A Comparative Analysis', https://satoriproject.eu/work_packages/comparative-analysis-of-ethics-assessment-practices/, (accessed on 22 February 2023).

MAGNETO, 'Fighting Against Crime and Terrorism', http://magneto-h2020.eu/, (accessed on 22 February 2023).

The Danish Institute for Human Rights, 'Human rights impact assessment guidance and toolbox', 2020, https://www.humanrights.dk/tools/human-rights-impact-assessment-guidance-toolbox, (accessed on 27 February 2023).