

Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights

ANDREA COSENTINI¹; ORESTE POLLICINO²; GIOVANNI DE GREGORIO³; ANDREA ERMELLINO¹;
DARIO FONTANELLA¹; NICOLE INVERARDI¹; FEDERICA PAOLUCCI²; ILARIA GIUSEPPINA PENCO¹;
DANIELE REGOLI¹; SILVIA TESSARO TRAPANI¹.

¹Intesa Sanpaolo S.p.A. ²Bocconi University ³Católica Global School of Law

Table of content: Executive summary. – 1. Introduction. – 2. The European Constitutional Framework and Requirements for FRIA; 2.1 The FRIA in the AI Act; 2.2 GenAI, FRIA and Risk Assessment. – 3. From Theory to Practice: A Proposal for Implementing the FRIA Framework; 3.1 The Questionnaire; 3.2 The Matrix; 3.3 FRIAct: Integration of Questionnaire and Matrix; 3.4 FRIAct Lifecycle. – 4. LoanLens: a testing ground use case; 4.1 Overview of the AI systems; 4.2 The application of the FRIAct. – 5. Conclusion – Acknowledgements – Appendix 1: Questionnaire for LoanLens. – Appendix 2: Questions proposal for the Matrix (*vademecum*).

Keywords: Risk Assessment – Fundamental Rights – Artificial Intelligence Act – FRIA – European Charter – Machine Learning - General Purpose AI Systems

Executive summary

This paper introduces a structured and comprehensive framework for conducting a Fundamental Rights Impact Assessment (FRIA), tailored specifically to address the challenges posed by high-risk Artificial Intelligence (AI) systems, including Generative AI (GenAI). The study is grounded in the obligations set forth by Regulation (EU) 2024/1689, known as the Artificial Intelligence Act (AI Act), which emphasises the necessity of assessing and mitigating risks to fundamental rights. The FRIA bridges the gap between regulatory obligations and practical implementation by providing a structured approach to assess and mitigate risks to fundamental rights.

This paper situates the FRIA as a critical tool mandated by the AI Act for certain high-risk AI systems and proposes a framework to conduct this assessment named FRIAct (Fundamental Rights Impact Assessment AI Act). The analysis proposed in this work positions between the theoretical commitments to fundamental rights protection and their practical operationalisation, providing deployers of AI systems with a replicable approach that aligns with European constitutional values and legal requirements. This approach also applies to GenAI systems, which often involve complex and large-scale implications for individuals and society.

The first part of the paper situates the FRIA within the broader European legal and constitutional framework, emphasising the foundational role of the Charter of Fundamental Rights of the European Union (CFREU). It explores how fundamental rights enshrined in the CFREU shape the obligations of the AI Act. Notably, the Regulation underscores the importance of ensuring that high-risk AI systems, including GenAI, comply with European constitutional values. As a matter of fact, GenAI systems have emerged as transformative technologies capable of reshaping industries, yet they also pose significant challenges. These include risks to privacy, data protection, access to effective remedies, and human dignity, as well as systemic concerns such as misinformation and the erosion of trust in automated systems. The AI Act explicitly extends its regulatory scope to include GenAI, introducing specific safeguards for models with systemic risks. These provisions address the need for enhanced transparency, robust oversight, and proactive risk mitigation in the design, development, and deployment of GenAI systems.

The second part introduces the FRIAct based on a two-pronged approach integrating qualitative and quantitative tools for assessing risks to fundamental rights. The qualitative tool, referred to as the Questionnaire, is designed to gather contextual and operational insights of the AI system,

including its purpose, affected populations, technical characteristics, and the broader societal and ethical implications of its deployment. It evaluates risks producing a Questionnaire Risk Indicator (QRI) that serves as the foundation for further analysis. The Matrix, on the other hand, adds an assessment specifically designed to produce a quantitative output: this purpose is achieved by systematically mapping potential qualitative impacts on specific rights and then attributing a quantitative score to those impacts – this is why we also refer to the Matrix’ outcomes as *semi-quantitative* ones. It evaluates risks based on two key dimensions – Severity and Probability of Occurrence – and calculates Impact Significance (IS) scores for each right as outlined in the CFREU. The FRIAct incorporates both the Questionnaire and the Matrix to generate FRIAct Scores, a final output that quantifies the system’s overall risk to each fundamental right. Both Questionnaire and Matrix are relevant in the phase of pre-deployment and monitoring. This design ensures flexibility, enabling its application across diverse AI systems and use cases, while maintaining alignment with the AI Act’s regulatory requirements.

The final part of the paper applies the FRIAct framework to a practical use case: LoanLens, a high-risk AI system designed for credit scoring of natural persons. LoanLens integrates a traditional Machine Learning System (MLS) with a Generative AI-powered Decision Support System (DSS), creating a hybrid approach that combines structured and unstructured data processing. This section demonstrates how the FRIAct evaluates the system’s risks, including transparency, fairness, privacy, and human oversight during the decision-making process. The Questionnaire identifies the system’s context, purpose, and operational risks, while the Matrix quantifies its impact on specific fundamental rights, such as privacy, non-discrimination, and human dignity. The case study underscores the importance of robust human oversight, as mandated by Article 14 of the AI Act, and the need for continuous monitoring to address evolving risks throughout the system’s lifecycle.

The paper concludes by arguing that the FRIAct framework not only fulfils the compliance requirements of the AI Act but also sets a benchmark for ethical and accountable AI deployment. The FRIAct framework represents a critical step toward embedding fundamental rights at the core of AI. The proposed approach highlights the necessity of collaboration between regulators, AI providers, and deployers to ensure that AI systems not only comply with legal standards but also uphold the societal values enshrined in the CFREU. By providing deployers with a structured approach to assess and manage risks, this framework operationalises fundamental rights protection in a way that is practical, replicable, and adaptable to diverse AI systems.

1. Introduction

The evolution of AI systems, including those embedding GenAI models, have been shaping the transformation of various sectors,¹ from healthcare and finance to social media and education.² As these systems become more pervasive, they bring about substantial benefits, including enhanced efficiency, personalised user experiences, and new creative possibilities. However, this transformative potential is accompanied by significant risks to fundamental rights such as privacy, freedom of expression, equality, and human dignity.³ These risks are not always immediately visible and often involve complex, large-scale repercussions that can affect both individuals and communities. For instance, GenAI models have been linked to issues such as copyright protection,⁴ algorithmic bias, data privacy breaches, misinformation, and surveillance concerns, each with far-reaching implications for fundamental rights.⁵

¹ Generative Artificial Intelligence (GenAI) refers to advanced AI systems that can create new content—such as text, images, audio, or video – based on vast datasets and complex algorithms. Unlike traditional AI systems, which typically perform predefined tasks, GenAI models are capable of producing original content by learning patterns from large datasets, often requiring minimal human intervention. These models are increasingly applied in various domains, including content creation, design, and predictive analysis, and have raised specific concerns around privacy, data security, and intellectual property. Mindy Nunez Duffourc Kollnig Sara Gerke & Konrad, ‘Privacy of Personal Data in the Generative AI Data Lifecycle’ (*NYU Journal of Intellectual Property & Entertainment Law*, 8 July 2024) <<https://jipel.law.nyu.edu/privacy-of-personal-data-in-the-generative-ai-data-lifecycle/>> accessed 4 November 2024; Francesco Corea, ‘AI Knowledge Map: How to Classify AI Technologies’ in Francesco Corea (ed), *An Introduction to Data: Everything You Need to Know About AI, Big Data and Data Science* (Springer International Publishing 2019) <https://doi.org/10.1007/978-3-030-04468-8_4> accessed 10 May 2022; Sandeep Singh Sengar and others, ‘Generative Artificial Intelligence: A Systematic Review and Applications’ (arXiv, 17 May 2024) <<http://arxiv.org/abs/2405.11029>> accessed 4 November 2024. For a detailed exploration of GenAI’s implications for personal data privacy, European Data protection Board (EDPB), ‘Report of the work undertaken by the ChatGPT Taskforce’, 23 May 2024; Taner Kuru, ‘Lawfulness of the Mass Processing of Publicly Accessible Online Data to Train Large Language Models’ [2024] *International Data Privacy Law* ipae013.

² Giovanni De Gregorio, ‘The Normative Power of Artificial Intelligence’ (2023) 30 *Indiana Law Journal*; Francesco Paolo Levantino, ‘Generative and AI-Powered Oracles: “What Will They Say about You?”’ (2023) 51 *Computer Law & Security Review* 105898; Natali Helberger and Nicholas Diakopoulos, ‘ChatGPT and the AI Act’ (2023) 12 *Internet Policy Review* <<https://policyreview.info/essay/chatgpt-and-ai-act>> accessed 25 May 2023; Claudio Novelli and others, ‘Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity’ (arXiv, 15 March 2024) <<http://arxiv.org/abs/2401.07348>> accessed 9 October 2024.

³ Georgios Feretzakis and Vassilios S Verykios, ‘Trustworthy AI: Securing Sensitive Data in Large Language Models’ (arXiv, 26 September 2024) <<http://arxiv.org/abs/2409.18222>> accessed 4 November 2024.

⁴ Uri Y Hachohen and Niva Elkin-Koren, ‘Copyright Regenerated: Harnessing GenAI to Measure Originality and Copyright Scope’ (3 August 2023) <<https://papers.ssrn.com/abstract=4530717>> accessed 26 February 2024.

⁵ Oreste Pollicino, Marco Fasciglione, Giovanni De Gregorio, Federica Paolucci ‘Compliance through Assessing Fundamental Rights: Insights at the Intersections of the European AI Act and the Corporate Sustainability Due Diligence Directive’ (*MediaLaws*, 30 July 2024) <<https://www.medialaws.eu/compliance-through-assessing-fundamental-rights-insights-at-the-intersections-of-the-european-ai-act-and-the-corporate-sustainability-due-diligence-directive/>> accessed 4 November 2024.

Against these challenges,⁶ the European Union has introduced a horizontal and comprehensive legal framework by enacting the AI Act.⁷ This regulation underlines the European intent to consider not only the development of cutting-edge technologies but also safeguarding “health, safety and fundamental rights”.⁸ As a matter of fact, in its first recitals, the AI Act primarily refers to European values,⁹ and the need to respect human dignity, freedom, equality, democracy, the rule of law and fundamental rights, as framed in Art. 2 TUE, and enshrined in the CFREU.

This framework, oriented to the protection of fundamental rights and democratic values, inspires the entire legal framework of the AI Act as driven by a risk-based approach.¹⁰ The regulation identifies different areas of risks corresponding to specific obligations moving from prohibited practices,¹¹ to systems with high and minimal risks related to the impact on fundamental rights. These obligations also extend to GenAI models, which have been introduced within the scope of the regulation during the political negotiation,¹² as a reaction to the spread of AI applications based on these models and the consequent rising attention of European and national institutions on the

⁶ European Commission, White Paper on Artificial Intelligence: a European approach to excellence and trust, 19 February 2020.

⁷ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

⁸ AI Act, Rec. 1-3.

⁹ Huw Roberts and others, ‘Safeguarding European Values with Digital Sovereignty: An Analysis of Statements and Policies’ (2021) 10 Internet Policy Review <<https://policyreview.info/articles/analysis/safeguarding-european-values-digital-sovereignty-analysis-statements-and-policies>> accessed 26 July 2022; Kim Lane Scheppele, Dmitry Vladimirovich Kochenov and Barbara Grabowska-Moroz, ‘EU Values Are Law, after All: Enforcing EU Values through Systemic Infringement Actions by the European Commission and the Member States of the European Union’ (2020) 39 Yearbook of European Law 3; Anu Bradford, ‘Europe’s Digital Constitution’ [2023] Verfassungsblog <<https://verfassungsblog.de/europes-digital-constitution/>> accessed 12 February 2024.

¹⁰ AI Act, Rec. 27. Claudio Novelli and others, ‘AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act’ (2024) 3 Digital Society 13; Pietro Dunn and Giovanni De Gregorio, ‘The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age’ (2022) 59 Common Market Law Review <<https://kluwerlawonline.com/journalarticle/Common+Market+Law+Review/59.2/COLA2022032>> accessed 20 May 2022.

¹¹ Ibid, Art. 5 AI Act.

¹² Nathalie A Smuha and Karen Yeung, ‘The European Union’s AI Act: Beyond Motherhood and Apple Pie?’ (24 June 2024) <<https://papers.ssrn.com/abstract=4874852>> accessed 14 August 2024.

matter.¹³ The AI Act specifically deals with GenAI models by introducing specific safeguards, particularly to those models which fall within the category of systemic risk.¹⁴

Among the key requirements expressing this constitutional orientation, the AI Act introduces the FRIA. This step requires assessing the risks for fundamental rights coming from the deployment of AI systems, thus increasing the accountability of public and private actors using these technologies. This obligation also includes GenAI models, which, if integrated into AI systems, aim at extending the entire architecture of fundamental rights protection to GenAI systems. By means of regulation, the FRIA is an evaluative process specifically designed for certain systems classified as high-risk by the AI Act.¹⁵ This requirement mandates that organisations deploying such systems undertake a comprehensive assessment to identify, measure, and mitigate potential negative impacts on fundamental rights. The FRIA is intended to prevent abuses and unintended consequences, fostering responsible AI usage that respects fundamental rights at all stages of development and deployment.¹⁶ By introducing FRIA, the AI Act establishes a framework for transparency, accountability, and ethical practices, which are essential for building public trust in AI technologies.

However, despite the relevance of this instrument, there are still no solid approaches to developing a FRIA for AI systems, including those systems based on GenAI models.¹⁷ Although

¹³ Italy's SA temporarily banned OpenAI's ChatGPT due to concerns over the unlawful collection and processing of personal data, and the lack of adequate mechanisms to prevent minors from accessing the platform. The case raised questions about how GenAI systems collect, store, and use personal information, especially when operating at such large scales. The Garante ordered OpenAI to implement stronger privacy safeguards and transparency measures, including age verification tools and greater user control over their data. The incident attracted widespread attention, marking the first regulatory action taken against a major GenAI system in Europe. See Italian SA, Measure of 30 March 2023, Register of Measures, No. 112 of 30 March 2023.

¹⁴ Specifically, the AI Act focuses on 'general purpose AI systems', defined by Art. 3, para. 1, no. 63) as "means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market".

¹⁵ AI Act, Rec. 96, "The aim of the fundamental rights impact assessment is for the deployer to identify the specific risks to the rights of individuals or groups of individuals likely to be affected, identify measures to be taken in the case of a materialisation of those risks. The impact assessment should be performed prior to deploying the high-risk AI system and should be updated when the deployer considers that any of the relevant factors have changed".

¹⁶ Alessandro Mantelero, 'The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, Legal Obligations and Key Elements for a Model Template' <<https://iris.polito.it/handle/11583/2991821>> accessed 4 September 2024.

¹⁷ In general, Art. 27 of the AI Act gives the competence to develop the template to the AI Office. Efforts towards the development of an explicit methodology made in recent years by different public institutions, including the government of the Netherlands, and the Danish Institute for Human Rights, as well as the Canadian Government. See Government of the Netherlands, 'Fundamental Rights and Algorithms Impact Assessment (FRAIA)', 31 July 2021; Danish Institute for Human Rights, 'Human Rights Impact Assessment and Toolbox', 2020; Government of Canada, 'Algorithmic Impact Assessment tool', <<https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>>.

research has focused on studying the possible adverse risks caused by such technologies,¹⁸ a generally adopted methodology for assessing the impact of AI systems on fundamental rights is still lacking.¹⁹ Addressing this gap, this paper proposes a structured FRIA framework that organisations can apply to assess the impact of the use of AI systems on fundamental rights. The proposed approach consists of two main components:

- **Questionnaire:** The open-ended questionnaire gathers context-specific information, examining the unique technical and operational characteristics of the AI system in question. By identifying possible risks inherent in the system's design and application, this analysis serves as the foundational step in assessing its impact on fundamental rights.
- **Matrix:** The matrix component assigns scores to various rights based on the assessed risk factors, facilitating a numerical assessment of threats to fundamental rights. This matrix incorporates indicators such as the likelihood of adverse outcomes, population exposure, and the severity of impact on each right. In cases where numerical indicators are not feasible, an ordinal scale (e.g., low, medium, high) is employed to ensure a structured evaluation of risks.

Together, these components establish a replicable approach for assessment that enables deployers to systematically evaluate and manage the risks posed by AI systems. This work aims to propose a practical and operational framework to obtain a concrete and reproducible approach to develop the FRIA for high-risk AI systems introduced by the AI Act. By integrating theoretical principles with practical applications, this work seeks to advance responsible AI governance, enabling deployers to navigate the complex interplay between innovation and fundamental rights in the digital age. This

¹⁸ European Commission, 'Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe', 2021; European Union Agency for Fundamental Rights, 'Getting the future right – Artificial intelligence and fundamental rights' (2020); European Agency for Fundamental Rights, 'Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights', 11 June 2019.

¹⁹ Some proposals have been developed. Allow the referral to a precedent work curated by Samuele Bertaina and others, 'Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of AI Act' (2025) 56 Computer Law & Security Review 106101. In this paper, a first methodology for conducting the FRIA has been proposed to underline the importance of accountability and risk based approach. This policy paper represents indeed an evolution of the methodology, which builds a more comprehensive understanding, including also GenAI systems.

It is worth mentioning that has been recently published by the Catalan Data Protection Authority a report which develops an assessment of the impact on fundamental rights in the use of artificial intelligence (called AIDF), in line with what is established in the new Artificial Intelligence Regulation. See DPD, 'FRIA model: Guide and use cases FRIA methodology for AI design and development', 2025.

paper thus serves as a crucial resource for policymakers, AI developers, and stakeholders seeking to integrate a rights-centred approach into the AI lifecycle, promoting the responsible deployment of AI in high-stakes environments.

The assessment procedure aims to provide numerical values, made up of the different evaluation phases, which make it possible to assess the risks coming from the deployments of AI systems. These final risk scores are achieved through the combination of the risk scores resulting from two factors: 1) the completion of a Questionnaire that investigates contextual, processual, and technical aspects of the AI system; 2) the development of a semi-quantitative Matrix assessing possible impacts on fundamental rights.

The first section situates the FRIA within the European legal and constitutional framework, emphasising the pivotal role of the CFREU and the foundational principles of European digital constitutionalism. This section underscores the legal and ethical imperatives for ensuring that AI systems respect and protect fundamental rights such as privacy, non-discrimination, and human dignity. The second section introduces the model for implementing the FRIAct, combining qualitative and quantitative tools. It details a two-pronged approach: a Questionnaire that systematically gathers contextual and technical information to identify potential risks, and a Matrix that evaluates the severity and likelihood of impacts on specific fundamental rights. Together, these tools provide a replicable and adaptable framework for assessing risks and mitigating potential harms resulting from the use of AI systems. The third section illustrates the practical application of the methodology through a case study of a high-risk AI system – LoanLens, a credit-scoring tool combining Machine Learning and GenAI technologies. This example demonstrates the framework's capacity to identify vulnerabilities, quantify risks, and guide mitigation strategies in real-world contexts, ensuring compliance with the AI Act and fostering trust in AI deployment.

2. The European Constitutional Framework and Requirements for FRIA

The expansion of the constitutional narrative in the field of AI is part of a broader trend that started before the spread of AI systems. Since the launch of the Digital Single Market Strategy in 2015, the European Union has moved from a framework dominated by a narrative of digital liberalism to a framework of digital constitutionalism characterised by a larger attention on the

protection of rights and freedoms.²⁰ The objective is to ensure that the logic of market freedoms does not override reasons of public interest, such as the protection of health, safety, and fundamental rights. At the same time, the European digital constitutional identity has also been based on the need to ensure fundamental freedoms and competition, which have played a foundation role in the EU economic integration process since the beginning, thus making economic freedoms critical to creating a market for AI in Europe.

The obligation to conduct the FRIA introduced by the AI Act falls within this broader framework which has led public and private institutions to increase their accountability in the digital age. Among different trends, Human Rights Impact Assessments (HRIA) have become essential for integrating human rights protections into business processes, following the UN Guiding Principles on Business and Human Rights.²¹ Likewise, at the international level, the Council of Europe, through initiatives of the Committee on Artificial Intelligence (CAI), elaborated the Human Rights Democracy and Rule of Law Impact Assessment (HUDERIA),²² which aimed to establish international AI standards focusing on risk identification, impact assessment, governance evaluation, and continuous mitigation to protect human rights.²³ In the elaboration of the AI Act, precisely under the impulse of the European Parliament, the EU adopted the FRIA as its own impact assessment method, also taking inspiration from the Data Protection Impact Assessment (DPIA) in the General Data Protection Regulation (GDPR).²⁴

The FRIA aims to identify risks to safeguard fundamental rights from high-risk AI systems potential impacts.²⁵ It mandates comprehensive pre-deployment assessments to prevent adverse

²⁰ Oreste Pollicino, *Judicial Protection of Fundamental Rights on the Internet: A Road Towards Digital Constitutionalism?* (Bloomsbury Publishing 2021). Giovanni De Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society* (Cambridge University Press 2022).

²¹ United Nations, 'Guiding Principles on Business and Human Rights in the technology sector', 2011.

²² CAI, Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law (HUDERIA Methodology) 28 November 2024 <https://rm.coe.int/cai-2024-16rev2-methodology-for-the-risk-and-impact-assessment-of-arti/1680b2a09f>

²³ For an analysis of the CoE attempts to establish common regulatory principles for AI, see Francesco Paolo Levantino and Federica Paolucci, 'Advancing the Protection of Fundamental Rights Through AI Regulation: How the EU and the Council of Europe Are Shaping the Future' (27 June 2024) SSRN <<https://papers.ssrn.com/abstract=4881656>>.

²⁴ 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) (Text with EEA relevance).

²⁵ AI Act, Rec. 96: "the impact assessment should identify the deployer's relevant processes in which the high-risk AI system will be used in line with its intended purpose, and should include a description of the period of time and frequency in which the system is intended to be used as well as of specific categories of natural persons and groups who are likely to be affected in the specific context of use. The assessment should also include the identification of specific risks of harm

impacts on fundamental rights, emphasising accountability, transparency, and ethical AI use. Given its strong connection with the European constitutional framework, particularly the protection of fundamental rights, it is crucial to connect the FRIA with the constitutional safeguards towards rights and freedoms in Europe as enshrined in the CFREU.

On these premises, the AI Act refers to European values and the FRIA, it recalls a series of elements related to the application and interpretation of, at least, the CFREU. As introduced in 2000 and then becoming legally binding since the Treaty of Lisbon in 2009, it preserves a wide array of civil, political, economic, and social rights, consolidating the rights derived from the constitutional traditions and international obligations common to the Member States. Nonetheless, from the outset, the European integration project was deeply rooted in the collective vision of shared values such as human dignity, democracy, the rule of law, and respect for human rights.²⁶ These values, later codified in the Treaties,²⁷ were envisioned as the foundation for unifying a continent fractured by war and conflict, guiding the Union's evolution toward safeguarding both individual and collective aspirations.

It is worth to be recalled that the fundamental rights protected by the Charter are not absolute.²⁸ The lack of a strict hierarchy opens to the balancing of different conflicting constitutional interests which “must be considered in relation to their function in society”.²⁹ The process of balancing is primarily connected with the principle of proportionality. According to Article 52 of the Charter, its provisions must be interpreted in a way that does not restrict or adversely affect the essence of rights. Specifically, the CJEU specified that any restriction to the enjoyment of such rights must not constitute an unreasonable, disproportionate and intolerable infringement of them. For example, in the *Schrems* case, the CJEU invalidated an EU decision because it compromised the essence of the

likely to have an impact on the fundamental rights of those persons or groups. While performing this assessment, the deployer should take into account information relevant to a proper assessment of the impact, including but not limited to the information given by the provider of the high-risk AI system in the instructions for use. In light of the risks identified, deployers should determine measures to be taken in the case of a materialisation of those risks, including for example governance arrangements in that specific context of use, such as arrangements for human oversight according to the instructions of use or, complaint handling and redress procedures, as they could be instrumental in mitigating risks to fundamental rights in concrete use-cases.” See, also, Mantelero (n 16).

²⁶ EU Fact Sheet, ‘The protection of EU values’, 2025.

²⁷ Scheppele, Kochenov and Grabowska-Moroz (n 9).

²⁸ The Charter does not explicitly identify the rights that are absolute. Based on the Charter explanations, the ECHR and the case law of the European courts, it is submitted that human dignity (Article 1 of the Charter) together with the prohibition of torture and inhuman or degrading treatment or punishment (Article 4 of the Charter), are to be considered absolute rights.

²⁹ C-311/18, *Data Protection Commissioner v Facebook Ireland Ltd, Maximillian Schrems*, ECLI:EU:C:2020:559, para. 172.

right to privacy and the right to judicial protection, setting a precedent that any measure infringing on the essence of a fundamental right is automatically invalid.³⁰ Thus, the principle of proportionality and the concept of the essence of fundamental rights are critical in protecting these rights, ensuring that limitations are justified, appropriate, and necessary.

Therefore, fundamental rights have become essential in ensuring the legality of actions by public actors within the EU,³¹ and their scope has also extended to private actors. Specifically, while Article 51(1) of the Charter does not explicitly state that private individuals have to ensure the protection of fundamental rights established by the Charter, the CJEU has ruled that some rights in the Charter apply directly between individuals, according to the doctrine that recognises the horizontal application of such rights.³² For instance, the prohibition of discrimination (Article 21) has direct effects on disputes between individuals.³³ Thus, the prominent application of such principles found its way precisely in the digital realm, where the CJEU ruled the existence of horizontal protection of individual's data rights, including de-listing information, in Articles 7 and 8 of the Charter.³⁴

Even if aligned with this framework when referring to the protection of European values, however, the AI Act does not fully focus on fundamental rights. Thus, despite some provisions of the AI Act,³⁵ the Regulation contains few concrete references,³⁶ which do not provide a comprehensive framework to ensure systematic and adequate verification of fundamental rights violations by specific AI systems. This holds true since the AI Act responds more to a product reliability approach rather than a fundamental rights instrument,³⁷ as also testified by its legal basis.³⁸

³⁰ Case 311/18, *supra*.

³¹ Case 29/69, *Stauder*, EU:C:1969:57; Case 11/70, *Internationale Handelsgesellschaft*, EU:C:1970:114; Case 4/73, *Nold*, EU:C:1974:51.

³² C-176/12, *AMS*, 2014, para. 47; C-414/16, *Egenberger*, 2018, paras 76 and 78; C-569/16 and C-570/16, *Bauer*, 2018, para. 85; C-684/16, *Max-Planck-Gesellschaft zur Förderung der Wissenschaften*, 2018, paras 77–9; C-193/17, *Cresco*, 2019, 76.

³³ C-43/75, *Gabrielle Defrenne v Société anonyme belge de navigation aérienne Sabena*, 8 April 1976.

³⁴ Eleni Frantziou, *The Horizontal Effect of Fundamental Rights in the European Union: A Constitutional Analysis* (OUP, 2019); C-92/09 and C-93/09, *Volker und Markus Schecke and Eifert*, 2010, paras 65–86.

³⁵ E.g., Recital 1 emphasizes the regulation's aim to promote human-centric, trustworthy AI while ensuring high protection levels for health, safety, and fundamental rights, including democracy, the rule of law, and environmental protection; Recital 5 recognizes the potential risks AI poses to public interests and fundamental rights.

³⁶ Such as in Articles 13.3(b)(ii) on transparency and user information, and 14.2 on human oversight.

³⁷ Marco Almada and Nicolas Petit, 'The EU AI Act: a medley of product safety and fundamental rights?', *EUI, RS.C, Working Paper*, 2023/59, <https://hdl.handle.net/1814/75982>.

³⁸ Article 114 TFEU. For a critique on the extensive use of this legal base by the EU for basically any rule that applies to the digital sector, see Erik Longo, 'Grounding Media Freedom in the EU: The Legal Basis of the EMFA' (2025) 7 *Rivista italiana di informatica e diritto* 14.

Particularly, the lack of individuals' rights and the focus mainly on the categories of deployer and provider symbolised the market approach rather than a fundamental rights-driven one.³⁹

In this sense, the FRIA can be considered a compromise to reconcile the internal market dimension with the protection of fundamental rights. The FRIA can be considered a bridge between the realm of AI systems and the protection of fundamental rights based on the principles enshrined in the Charter, thus guiding the interpretation and operationalisation of fundamental rights within the framework of the AI Act. This interconnectedness underscores the importance of a robust compliance system that respects and upholds fundamental rights as a core component of the AI regulatory landscape.

2.1 The FRIA in the AI Act

Article 27 is a crucial provision aimed at ensuring that specific high-risk AI systems are used in ways that respect fundamental rights. The introduction of FRIA in the AI Act stems from the need to balance innovation with risk management, particularly concerning fundamental rights. It represents a significant advancement in the regulatory landscape and sets a precedent for how AI should be managed and monitored, also building on the Ethics Guidelines for Trustworthy AI and on the White Paper on AI, in which it is acknowledged that AI can lead to breaches of fundamental rights, such as freedom of expression, assembly, human dignity, non-discrimination, and privacy. The requirement for a thorough FRIA by both public and selected private entities is a proactive measure to prevent potential abuses and unintended negative consequences of AI deployment.⁴⁰

The AI Act provides few details on how this assessment should be conducted. Precisely, it specifies that it should include usage descriptions, affected groups, potential risks, human oversight measures, and risk mitigation steps.⁴¹ Once the deployer performs the assessment, this should be

³⁹ On the remedies and their efficacy in the digital sphere, De Gregorio, Giovanni and Demkova, Simona, *The Constitutional Right to an Effective Remedy in the Digital Age: A Perspective from Europe* (January 31, 2024). In van Oirsouw, Ch., de Poorter, J.; Leijten, I.; van der Schyff, G.; Stremmer, M.; de Visser, M. (eds), *European Yearbook of Constitutional Law* (forthcoming, 2024).

⁴⁰ As designated by Art. 27 para. 1, the referral is to "private entities providing public services", as well as, under Annex III, Art. 5, lett. b and c, "AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems used for the purpose of detecting financial fraud", and "AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance".

⁴¹ Art. 27, para. 1, lett a-f), and Rec. 96: "The impact assessment should identify the deployer's relevant processes in which the high-risk AI system will be used in line with its intended purpose, and should include a description of the period of time and frequency in which the system is intended to be used as well as of specific categories of natural persons and groups who are likely to be affected in the specific context of use. The assessment should also include the identification of specific risks of harm likely to have an impact on the fundamental rights of those persons or groups".

reported to the market surveillance authority and updated if necessary.⁴² Since the AI Act does not provide thorough guidance on how to conduct the assessment, as happened for the DPIA, it will be crucial to consult the guidelines and the templates elaborated by the European Authorities: in the case of the FRIA, the AI Act, assign this competence to the AI Office.⁴³ Hence, the provision of a standardised template for these assessments by the AI Office is a strategic move to ensure consistency and comprehensiveness in the evaluations. Furthermore, the obligation to report these assessments to the market surveillance authority ensures that there is oversight and that any identified risks are systematically addressed. This approach reflects a broader commitment to integrating ethical considerations into the development and deployment of AI, aligning with the EU's broader agenda of promoting trustworthy AI.⁴⁴

However, the haste with which the FRIA was introduced,⁴⁵ its real connection with the rest of the Regulation, which, as mentioned above, has its core in the product safety-based regulation, and the way in which designated private and public actors will have to concretely carry it out are all questions that will have to be answered a year from the entry into force of the AI Act.⁴⁶ With the aim of working on compliance before the norm becomes effectively applicable, it is crucial to identify a method, which is the scope of this paper, and to analyse the doubts in the interpretation of the norm.

The AI Act, being first and foremost a regulation oriented to the specific use of a product, better AI, focuses indeed, but not exclusively, on who utilises the product. Providers are tasked with ensuring compliance with essential requirements and undergoing relevant conformity assessments,⁴⁷ while certain deployers must adhere to provider instructions and conduct a FRIA prior to deploying high-risk AI systems.⁴⁸ Specifically, this obligation applies to three key categories of actors: 1) deployers governed by public law, such as public authorities, like law enforcement using face recognition; 2) private operators providing public services, including education, healthcare, social services, housing, and justice administration;⁴⁹ 3) operators deploying high-risk AI systems

⁴² Art. 27, para. 3.

⁴³ Art. 27, para. 5.

⁴⁴ Rec. 1 of the AI Act, and Explanatory Memorandum of the Proposal, published by the EU Commission in April 2021.

⁴⁵ The FRIA is the result of the EP version of the Regulation, published in June 2023.

⁴⁶ The rules relating to 'high risk' systems, as is the case with Art. 27, are set to apply one year after the regulation comes into force: hence, by 1 August 2025.

⁴⁷ AI Act, Art. 43.

⁴⁸ AI Act, Rec. 96.

⁴⁹ Note that even though art. 27 para. 1 is vague in defining this category, Rec. 96 specifies that it is mandated to perform a FRIA also, "services important for individuals that are of public nature may also be provided by private entities. Private entities providing such public services are linked to tasks in the public interest such as in the areas of education, healthcare, social services, housing, administration of justice".

intended for creditworthiness evaluation, credit scoring, or risk assessment and pricing in life and health insurance contexts, as per Annex III AI Act, points 5, lett. b) and c).⁵⁰ By mandating FRIAs solely for these actors, the AI Act emphasises the importance of anticipating and mitigating risks to fundamental rights before deploying high-risk AI technologies.

The FRIA must be performed prior to the initial deployment of the high-risk AI system and should be continued during the whole AI life cycle.⁵¹ This entails a thorough evaluation of the system's potential to adversely affect fundamental rights. For this reason, deployers may rely on previously conducted assessments, including those provided by the system's developer, provided the system's context of use remains unchanged.⁵² However, should any significant modifications arise during the system's use, such as changes in its purpose, affected populations, or identified risks, the deployer is obligated to update the assessment to reflect the new circumstances.⁵³

The assessment, whose completion should be notified to the relevant market surveillance authority, must comprehensively address the following aspects:

- System use and purpose: a detailed description of how the AI system will be integrated into the deployer's processes, ensuring alignment with its intended purpose.
- Operational parameters: an outline of the timeframe and frequency of the system's deployment.
- Affected individuals and groups: an identification of the categories of natural persons and communities likely to be impacted by the system, particularly marginalised or vulnerable groups.

⁵⁰ AI Act, Annex III, Art. 5 lett. b) and c), "5. Access to and enjoyment of essential private services and essential public services and benefits: [...] (b) AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems used for the purpose of detecting financial fraud; (c) AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance".

⁵¹ AI Act, Rec. 96, "Whilst risks related to AI systems can result from the way such systems are designed, risks can as well stem from how such AI systems are used. Deployers of high-risk AI system therefore play a critical role in ensuring that fundamental rights are protected, complementing the obligations of the provider when developing the AI system. Deployers are best placed to understand how the high-risk AI system will be used concretely and can therefore identify potential significant risks that were not foreseen in the development phase, due to a more precise knowledge of the context of use, the persons or groups of persons likely to be affected, including vulnerable groups."

⁵² AI Act, Art. 27 para. 2.

⁵³ Id.

- Risk analysis: a detailed evaluation of specific risks of harm that may arise from the system's use, with reference to information provided by the developer in compliance with Article 13 of the AI Act.
- Oversight and governance: a description of the human oversight measures that will be implemented, following the developer's guidelines.
- Mitigation measures: strategies to address the materialisation of identified risks, including internal governance mechanisms and accessible complaint procedures.

From the breakdown of the essential requirements, there emerges a potential information asymmetry between deployers and providers: much of the information hereby requested, such as the operational parameters must be shared by the provider with the deployer in order to make the assessment work. As noted by scholars, “providers possess key knowledge about system properties and technical limitations, impacting risk assessments during deployment”.⁵⁴ Therefore, a crucial element, even though not specifically listed in the AI Act formulation, is indeed that of a collaboration between the provider and the deployer.

Moreover, the success of these assessments relies heavily on the transparency and reliability of the information exchanged. Providers must supply accurate and comprehensive data regarding the design, functionality, and potential risks of the AI systems, enabling deployers to evaluate their societal and rights-based impacts on the concrete context of deployment of the technology. Furthermore, the AI Act does not mention the complex framework of fundamental rights that Member States need to take into account, being all of them also signatories of the European Convention of Human Rights (ECHR).⁵⁵ This aspect reflects the need for balancing fundamental rights in European constitutionalism, which does not only come from the constitutional approach of the Charter but also from the ECHR. Additionally, the absence of explicit mention of the criterion through which such rights should be balanced in their application might be a problematic element,⁵⁶

⁵⁴ Piergiorgio Chiara and Federico Galli, ‘Normative Considerations on Impact Assessments in EU Digital Policy’ [2024] MediaLaws <<https://www.medialaws.eu/rivista/normative-considerations-on-impact-assessments-in-eu-digital-policy/>> accessed 5 September 2024.

⁵⁵ Francesco Paolo Levantino and Federica Paolucci, ‘Advancing the Protection of Fundamental Rights Through AI Regulation: How the EU and the Council of Europe are Shaping the Future’, *SSRN*, 2024.

⁵⁶ For instance, the Member States of the EU are also signatories of the European Convention of Human Rights (ECHR) which might create slightly differences in the interpretation of the application of some provisions related to the protection of fundamental rights, as, for instance, stated in ECtHR 30 June 2005, No. 45036/98, *Bosphorus Hava Yolları Turizm ve Ticaret Anonim Şirketi v Ireland*.

especially since also private actors must conduct the assessment, as in the context of evaluating the ‘creditworthiness’ and the ‘risk assessment and pricing’ in case of insurance.

Likewise, the coordination between the FRIA and other assessments, such as the DPIA, is not fully addressed. This is not only a problem of consistency in legal terms but also a practical one in figuring out who should conduct both a FRIA and a DPIA as an obligation under the GDPR. Accordingly, the AI Act provides that the deployer must conduct the assessment before putting the system into service. Particularly, this ex-ante evaluation is the responsibility of deployers of high-risk systems in situations listed under Article 27 of the AI Act. Moreover, in case the deployer needs to perform a DPIA, the fundamental rights impact assessment referred to shall complement it.⁵⁷ Hence, the two compliance measures tend to overlap,⁵⁸ creating potential legal uncertainty between the two obligations. This aspect is particularly relevant not only in the cases when a deployer buys the service or the license from the provider, but also when the subject of the deployer and the provider coincide.

This is a significant and potential backlash of the AI Act for at least two reasons: first, the deployer may not participate in the design and development phases, which are carried out by the provider; second, the roles of the deployer and provider under the GDPR might be confused with respect to the different assessment methods. As a matter of fact, the obligation to perform these assessments often falls on the deployer, especially if classified as a ‘data controller’ under data protection law, making them responsible for conducting the DPIA as well. However, if a provider has significant control over an AI system, such as with foundation models, data protection authorities might classify the provider as a data controller or potentially a joint controller with the deployer. This classification carries additional responsibilities and obligations under data protection laws.⁵⁹

2.2 GenAI, FRIA and Risk Assessment

Together with the focus on creating a safe and fundamental rights-oriented market for AI, the AI Act has introduced, among general rules, also sectorial regulations for some uses or specific systems of AI. The latter is the case of GenAI systems, reflecting their growing importance and widespread

⁵⁷ AI Act, Art. 27(4).

⁵⁸ Federica Paolucci, ‘Shortcomings of the AI Act’, *Verfassungsblog*, 14 March 2024.

⁵⁹ *Id.*

use. Unlike other AI systems, GenAI systems are characterised by their broad applicability and the substantial impact they can have across various sectors. This has led to a distinct set of rules within the AI Act, aimed at addressing the unique challenges posed by these systems, as introduced by the European Parliament before the trilogue negotiations in June 2023.

GenAI systems are particularly disruptive because of the vast number of parameters used in their training and their versatile nature. For this reason, identifying and applying regulations to GenAI systems is very challenging. These systems can impact a wide range of societal functions, requiring comprehensive regulatory frameworks.

These complexities are also mirrored by the classification adopted by the AI Act. As a matter of fact, the regulation classifies AI systems into different risk levels. However, GenAI models are subject to specific rules outlined in Chapter V, separate from the general risk categorisation in Chapters II-IV. The AI Act aims to provide specific rules for generative AI models, particularly those posing systemic risks, which also apply when these models are integrated in an AI system. As clarified by Rec. 97, “this Regulation provides specific rules for general-purpose AI models and for general-purpose AI models that pose systemic risks, which should apply also when these models are integrated or form part of an AI system. It should be understood that the obligations for the providers of general-purpose AI models should apply once the general-purpose AI models are placed on the market”.

Hence, the rules on GenAI distinguish if the system is to be considered as posing a ‘systemic risk’ or not. Therefore, before continuing this analysis is to be clarified what we mean by systemic risk. Art. 3 no. 65) clarifies that “‘systemic risk’ means a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain”.

In this respect, GenAI models need to be integrated in AI systems in order to become specific-purpose systems and therefore classifiable as high-risk. The obligations for providers of GenAI models apply when these models are placed on the market and continue when integrated into an AI system.⁶⁰ These obligations are distinct from those for AI systems and do not apply when the model is used exclusively for internal processes that do not affect individual rights or provide third-party

⁶⁰ Rec. 97.

services. However, GenAI models with systemic risks are always subject to the AI Act's obligations due to their potential adverse effects.

However, this definition, highly criticised at a scholarly level,⁶¹ might create some ambiguities with the identification of the obligations specifically for GenAI and the ones set for high-risk systems. Since the AI Act provides a structured framework for overseeing GenAI systems, stressing the importance of specific rules for these versatile and potentially disruptive technologies, the Regulation ensures that GenAI models, when integrated into AI systems, are subject to comprehensive regulatory oversight, reflecting their significant impact on society.

When a GenAI model is to be considered as posing a 'systemic risk', in addition to the obligation of transparency set by Artt. 53 and 54, Art. 55 requests that the providers of such model should: 1) perform evaluations using standardized protocols, including adversarial testing to identify and mitigate systemic risks; 2) assess and mitigate potential systemic risks at the Union level stemming from the model's development, market placement, or usage; 3) document and report serious incidents and corrective measures to the AI Office and relevant national authorities without delay; ensure robust cybersecurity measures for both the AI model and its physical infrastructure.

Specifically, the conformity and mitigation assessment under the AI Act should be conducted to tackle any 'systemic risk'. To understand what the purpose of such analysis is, it is crucial to read Art. 55 para. 1 lett. b) together with the definition of 'systemic risk' provided by Art. 3. Under these lenses, it is possible to understand that conformity assessment under Art. 55 must regard "any actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain".⁶²

However, the FRIA is not a standalone instrument but needs to be complemented by the 'conformity assessment' conducted by the provider to address the effective risks and harms to various categories of people. When concretely possible, it would be ideal for the provider and the deployer to establish a dialogue throughout the AI value chain to evaluate and respond to the "measures to be taken in case of the materialisation of those risks, including the arrangements for

⁶¹ Philipp Hacker, 'What's Missing from the EU AI Act: Addressing the Four Key Challenges of Large Language Models', *VerfassungsBlog*, 13 December 2023, <https://verfassungsblog.de/whatsmissing-from-the-eu-ai-act/>; N.A.Smuha, K. Yeung, "The European Union's AI Act: beyond motherhood and apple pie?", (June 24, 2024), *SSRN*; P. Hacker, A. Engel, and M. Mauer, "Regulating ChatGPT and other large generative AI models", *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 2023.

⁶² See AI Act, Art. 3 no. 65. For a comparison, read A. Mantelero, "The Fundamental Rights Impact Assessment (FRIA) in the AI Act: roots, legal obligations and key elements for a model template", *Computer Science Law Review*, 2024.

internal governance and complaint mechanisms”.⁶³ Without significant collaboration between the provider and the deployer, it would be complex to have a comprehensive view of the risks potentially driven by GenAI. This is crucial not only due to the inherent complexity of these models but also to ensure that risks are managed effectively at both the model and application levels.

Before delving into the framework proposed for conducting a FRIA, it is crucial to situate this assessment within the broader landscape of risk evaluations established by the AI Act. Since several provisions under the AI Act address risk management, including Article 9’s Risk Management System (RMS),⁶⁴ and Article 43’s Conformity Assessment (CA),⁶⁵ it is critical to focus on the interplay between the FRIA and the other assessment obligations, as this integration ensures a cohesive approach to compliance.

Essentially, Article 9 establishes a RMS⁶⁶ as a core requirement for all high-risk AI systems, emphasising the identification and mitigation of risks to health, safety, and fundamental rights.⁶⁷ As mentioned, the FRIA is a targeted assessment specifically applied to certain high-risk AI systems deployed by public bodies, private entities providing public services, or operators involved in high-stakes applications, such as credit scoring or health and life insurance. In contrast, Article 9 applies universally to all high-risk AI systems, covering risks across health, safety, and fundamental rights during the design and development phases.⁶⁸

⁶³ AI Act, Article 27 (1) lett. f).

⁶⁴ Specifically, as the SRA, article 9 establishes a risk management system that should be performed by the provider. The definition of a risk management system is a process that “specifies how providers of high-risk AI systems must identify, assess and respond to risks”, cit. Risto Uuk and others, ‘A Taxonomy of Systemic Risks from General-Purpose AI’ (Social Science Research Network, 22 November 2024) <<https://papers.ssrn.com/abstract=5030173>> accessed 3 January 2025.

⁶⁵ Whereas article 43’s conformity assessment ensures compliance with the AI Act’s requirements through external validation “in order to ensure a high level of trustworthiness of high-risk AI systems”, as specified by Rec. 132.

⁶⁶ The definition of a risk management system is a process that “specifies how providers of high-risk AI systems must identify, assess and respond to risks”, cit. Jonas Schuett, ‘Risk Management in the Artificial Intelligence Act’ (2024) 15 European Journal of Risk Regulation 367.

⁶⁷ AI Act, Art. 9, para. 2: “. The risk management system shall be understood as a continuous iterative process planned and run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic review and updating. It shall comprise the following steps:

- (a) the identification and analysis of the known and the reasonably foreseeable risks that the high-risk AI system can pose to health, safety or fundamental rights when the high-risk AI system is used in accordance with its intended purpose;
- (b) the estimation and evaluation of the risks that may emerge when the high-risk AI system is used in accordance with its intended purpose, and under conditions of reasonably foreseeable misuse;
- (c) the evaluation of other risks possibly arising, based on the analysis of data gathered from the post-market monitoring system referred to in Article 72;
- (d) the adoption of appropriate and targeted risk management measures designed to address the risks identified pursuant to point (a).”

⁶⁸ AI Act, Rec. 65.

In terms of actors that are responsible for complying with the obligations, the FRIA places responsibility on deployers – entities deploying the AI system in real-world contexts. These deployers must consider risks specific to the operational setting, including impacts on marginalised groups. Conversely, RMS under Article 9 is the obligation of providers that incorporate risk identification and mitigation measures into the system’s design, adhering to essential requirements.⁶⁹

Furthermore, the AI Act imposes that providers of GenAI models perform SRA (Systemic Risk Assessment), as mentioned above, focused primarily on the GenAI model itself rather than its specific application.⁷⁰ This approach is criticised for addressing hypothetical risks that may never materialise rather than managing concrete risks associated with the AI’s actual deployment.⁷¹ Hence, this conformity assessment, tackling also the impact on fundamental rights, must be carried out when GenAI is used in high-risk AI systems or general-purpose models that pose a systemic risk. Hence, while the provider of a GenAI posing a systemic risk must complete a ‘conformity assessment’ under Article 55, it is important to note that, if the AI is applied for a purpose classified as high-risk under Annex III, the deployer will also be required to perform a FRIA. Thus, a double track of assessment would take place in which, on the one hand, the provider and, on the other hand, the deployer assess the even potential impact of GenAI on fundamental rights.

In a nutshell, the SRA requires providers to conduct adversarial testing, assess systemic risks at the Union level, ensure cybersecurity,⁷² and report serious incidents to the AI Office and national authorities. Unlike the FRIA, which is the responsibility of deployers to evaluate the operational context, the SRA places obligations on providers to mitigate risks embedded in the design and development of AI models. This division of responsibilities reflects the complementary roles of these

⁶⁹ AI Act, Art. 9, para. 5, specifies that identifying the essential requirements for performing the risk management system are “(a) elimination or reduction of risks identified and evaluated pursuant to paragraph 2 in as far as technically feasible through adequate design and development of the high-risk AI system; (b) where appropriate, implementation of adequate mitigation and control measures addressing risks that cannot be eliminated; (c) provision of information required pursuant to Article 13 and, where appropriate, training to deployers. With a view to eliminating or reducing risks related to the use of the high-risk AI system, due consideration shall be given to the technical knowledge, experience, education, the training to be expected by the deployer, and the presumable context in which the system is intended to be used”.

⁷⁰ AI Act, Rec. 114, “The providers of general-purpose AI models presenting systemic risks should be subject, in addition to the obligations provided for providers of general-purpose AI models, to obligations aimed at identifying and mitigating those risks and ensuring an adequate level of cybersecurity protection, regardless of whether it is provided as a standalone model or embedded in an AI system or a product. To achieve those objectives, this Regulation should require providers to perform the necessary model evaluations, in particular prior to its first placing on the market, including conducting and documenting adversarial testing of models, also, as appropriate, through internal or independent external testing”.

⁷¹ A. Mantelero, *id.*; P.Hacker, *id.*

⁷² AI Act, Rec. 115.

assessments. Furthermore, the FRIA emphasises context-specific evaluations of risks such as data protection breaches, unfair outputs, and human oversight failures. In contrast, the SRA addresses risks related to the performance of the model, requiring standardised evaluation protocols and incident reporting to mitigate vulnerabilities inherent in general-purpose AI.

Hence, the SRA under Article 55 of the AI Act complements the FRIA under Article 27 by addressing different aspects of AI governance. While the FRIA focuses on deployment-specific risks to fundamental rights in high-risk AI systems, the SRA targets systemic risks posed by general-purpose AI models with broad societal impact. Therefore, given the distinction explained above, there might be GenAI which are classified as high-risk under the AI Act, but that are also deemed to pose systemic risks, as underlined by Art. 3 no. 65 of the Regulation. In this case, the provider should perform the SRA, and the deployer will perform a FRIA. Otherwise, if the system is not falling under one of the categories listed by Annex III, only the SRA will be conducted by the provider.⁷³

Together, these assessments create a comprehensive regulatory framework, with the SRA ensuring robustness and reliability at the model level and the FRIA addressing deployment-specific risks to fundamental rights. Building on these considerations, the following grid provides a comparative overview of these mechanisms, bridging the legal and technical dimensions of AI governance, and highlighting their roles in a comprehensive regulatory framework that balances innovation with fundamental rights protection.

Table 1 – Confrontation of the AI Act's assessment.

<i>Aspect</i>	<i>FRIA (Fundamental Rights Impact Assessment)</i>	<i>Risk Management System (RMS)</i>	<i>Systemic Risk Assessment (SRA)</i>
Objective	Assess and mitigate the impacts of high-risk AI systems on fundamental rights in specific deployment contexts.	Manage broader risks to health, safety, and fundamental rights during the design and development phases of AI systems.	Assess and mitigate systemic risks of general-purpose AI (GPAI) models at the Union level, including cybersecurity and societal risks.
Legal basis	Article 27 of the AI Act, focusing on fundamental rights as per the CFREU.	Article 9 of the AI Act, requiring a risk management system for all high-risk AI systems.	Article 55 of the AI Act, targeting systemic risks of GPAI models with potential EU-wide impact.

⁷³ Risto Uuk and others, id.

Actors	Deployers of high-risk AI systems, including public bodies and private entities providing public services.	Providers of high-risk AI systems during the development and design phases.	Providers of general-purpose AI models identified as posing systemic risks.
Timing	Conducted pre-deployment and updated as needed during the lifecycle of the AI system.	Iterative and continuous throughout the AI system lifecycle.	Performed periodically to monitor and address systemic risks.
Methodology	Primarily qualitative; lacks a standardized and unified methodology.	Often based on harmonized standards but focuses on technical design and lifecycle risk management.	Broad evaluation of systemic risks, often lacking specific focus on individual fundamental rights.
Focus on fundamental rights	Directly addresses fundamental rights with a broad focus on CFREU rights but may lack operational depth.	Includes fundamental rights alongside broader risks such as safety and performance.	Addresses fundamental rights indirectly by focusing on systemic-level impacts.
Outcome	Structured report outlining risks to fundamental rights and mitigation measures, submitted to market surveillance authorities.	Ongoing risk management plan ensuring compliance with harmonized standards.	Recommendations to mitigate systemic risks, such as cybersecurity or societal-level harms.

The regulatory landscape of the AI Act reflects a layered approach to managing risks posed by AI systems, particularly high-risk applications. As examined above, each tool serves a specific purpose, from mitigating individual and societal risks to ensuring technical compliance and safeguarding democratic values. While RMS and the SRA primarily concern providers, who focus on the design and development phases, conversely, the FRIA centers on deployers, emphasising the evaluation of risks in the deployment and operational context.

The FRIAct framework integrates these roles into a cohesive dual-phase approach of pre-deployment and post-deployment, ensuring the FRIA remains an active tool throughout the AI system's lifecycle. Essentially, the proposed methodology is designed to align pre-deployment and post-deployment efforts effectively, ensuring a comprehensive evaluation of risks and their mitigation in both design and monitoring phases. This approach bridges the gap between provider-

driven development and deployer-focused implementation, particularly when these entities are distinct. In case the provider and deployer are the same, the process is streamlined internally; however, when separate, FRIAct offers a structured mechanism to align their responsibilities effectively. By fostering collaboration and clearly delineating responsibilities, FRIAct not only enhances compliance with the AI Act but also strengthens its core mission: embedding fundamental rights protection seamlessly across all stages of AI governance.

3. From Theory to Practice: A Proposal for Implementing the FRIA Framework

The FRIA framework, as conceptualised in the AI Act, is designed to systematically evaluate risks to fundamental rights posed by high-risk AI systems. Building on the theoretical foundations of the FRIA, it is necessary to move toward practical implementation. This framework, referred to as the Fundamental Rights Impact Assessment AI Act (FRIAct),⁷⁴ adopts a risk-based approach intended to proactively anticipate and mitigate adverse effects on fundamental rights before they occur by adopting a process divided in three main phases:

- Phase 1 – Risk Classification Assessment, Intended Purpose, and Contextual Analysis. Identification whether the system qualifies as high-risk and frames its intended use. This phase is crucial since it provides the foundation for the entire assessment by establishing the context necessary for justifying decisions made in subsequent phases.
- Phase 2 – AI System Mapping & Monitoring. Examination of technical aspects and production of a Questionnaire Risk Indicator (QRI) that measures risks. The questions in this phase are designed to produce the QRI, a continuous score ranging from 1 (lowest risk) to 10 (highest risk). Each specific answer in this phase is assigned a corresponding risk level, represented as an integer from 1 to 10.⁷⁵
- Phase 3 – Fundamental Rights Impact Evaluation. Synthesis of these findings by integrating the outputs of the Matrix⁷⁶ –Impact Significance (IS) scores – into the Questionnaire to calculate the FRIAct Scores, essentially one for each fundamental right analysed, and,

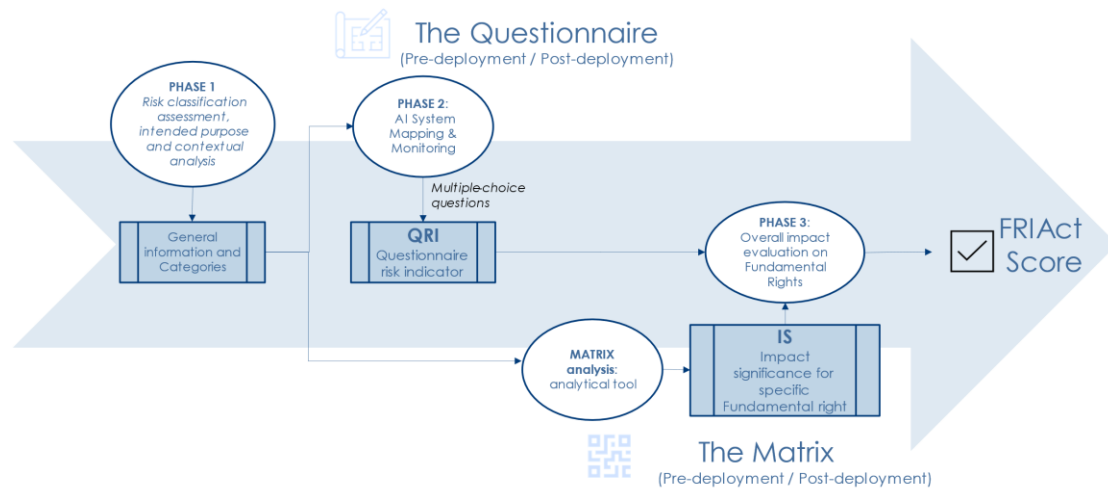
⁷⁴ Bertaina and others (n 19).

⁷⁵ For each section within Phase 2, a Section Risk Indicator is calculated as the average of the risk levels assigned to the answers within that section, resulting in a continuous score between 1 and 10. The ORI is then determined as the arithmetic average of all Section Risk Indicators across the Questionnaire. However, specific high-risk answers, identified as potential blockers, may override the calculated Section Risk Indicator and compromise the entire assessment.

⁷⁶ The combination of QRI with the IS scores will be explained in the following sections.

therefore, defining explicit mitigation measures based on the findings of the analysis, as requested by the AI Act.⁷⁷

Figure 1: A scheme of the assessment process, during which QRI and IS are calculated and then combined in the FRIAct Score. The process consists of two tools (Questionnaire and Matrix) and four Phases.



To achieve this objective, the FRIAct employs a two-pronged approach combining qualitative and quantitative tools: a Questionnaire and a Matrix. The Questionnaire is a qualitative tool structured to collect detailed contextual and technical information about the AI system. It serves as a foundation for understanding the system's purpose, design, and deployment environment, enabling a nuanced identification of risks and potential mitigations. Complementing this, the Matrix assesses the potential impacts on individual fundamental rights by calculating risk scores based on Severity and Probability of Occurrence dimensions, as described in detail in Section 4.2. Together, these tools ensure a comprehensive and reproducible evaluation, bridging theoretical principles with actionable insights for compliance and risk mitigation.

The methodology that has been implemented to perform such an assessment is tailored to adapt across the lifecycle of AI systems, encompassing two critical phases: pre-deployment and ost-deployment. The first assessment is conducted prior to system deployment, providing foundational insights for a 'Go/No-Go' decision. Whereas the post-deployment assessment, carried out

⁷⁷ AI act, para. 1, lett. e).

periodically, evaluates the evolving performance and impacts of the system, ensuring ongoing alignment with fundamental rights obligations.⁷⁸

The integration of the Questionnaire and Matrix ensures that FRIAct not only supports compliance with the AI Act but also advances best practices in responsible AI governance. By providing a practical and adaptable framework, FRIAct empowers to safeguard fundamental rights in the use of AI systems. The following sections detail the structure and application of these tools, followed by an in-depth analysis of the Questionnaire and followed by the Matrix, illustrating how they work in tandem to support an effective FRIA process.

3.1 The Questionnaire

The Questionnaire serves as the qualitative foundation of the FRIAct framework, systematically analysing the AI system's design, deployment, and operational context to assess potential impacts on fundamental rights. To ensure adaptability across different stages of the AI system lifecycle, such as the pre-deployment and post-deployment, a specific Questionnaire is carried out during the design stage, while a different Questionnaire is dedicated to assessing the AI system in production - offering a comprehensive assessment process that aligns with the AI Act's requirements. The Questionnaire is organised into three distinct phases, each addressing a critical aspect of the system's evaluation.

The first set of questions acknowledges whether the AI system qualifies as high-risk under Article 6 and Annex III of the AI Act and establishes the broader context of its deployment. Even if an AI system does not fall under the list included in Annex III of the AI Act, and, therefore, is not classified as high-risk, it can potentially impact natural persons significantly. In case the system is perceived to be potentially harming fundamental rights, then the deployer might consider performing the FRIAct in any case.

⁷⁸ A key challenge in developing this methodology was integrating diverse AI technologies, such as Machine Learning, Generative AI, and General-Purpose AI systems, into a cohesive framework. These technologies exhibit varied characteristics and interact differently with their operational environments, and, therefore, this diversity necessitates a flexible and adaptable approach, ensuring that FRIACT can address the wide spectrum of AI applications covered by the AI Act. By accommodating technological and operational differences, FRIACT provides a rational analysis of AI system impacts across varied use cases. For a distinction of the different characteristics of the technologies, see Giovanni Sartor, *L'intelligenza artificiale e il diritto* (Giappichelli 2022); Ethem Alpaydın, *Machine Learning* (MIT Press Essential Knowledge series 2016); Ziwei Liu and others, 'Generative Networks' in Stan Z Li, Anil K Jain and Jiankang Deng (eds), *Handbook of Face Recognition* (Springer International Publishing 2024) <https://doi.org/10.1007/978-3-031-43567-6_3> accessed 25 October 2024; Oskar J Gstrein, Noman Haleem and Andrej Zwitter, 'General-Purpose AI Regulation and the European Union AI Act' (2024) 13 Internet Policy Review <<https://policyreview.info/articles/analysis/general-purpose-ai-regulation-and-ai-act>> accessed 31 December 2024.

Therefore, the objectives of this first phase of the questionnaire are:

- Acknowledge whether the system falls under a certain high-risk category, initiating the obligation to conduct a FRIA.
- Frame the AI system within the scope of its project, detailing its operational purpose and potential societal impact.
- Map the lifecycle of the AI system, including key actors and affected persons.

Key questions of this phase include the understanding of the functionality of the system, the timeline and usage, the affected population, and the risk classification. This phase is crucial to frame the AI system within its lifecycle, identifying key stakeholders and contextual factors that inform technical assessments in later phases.

The second phase consists of several sections, each section containing multiple choice questions to which a risk level is assigned according to the answer selected. In this phase, the Questionnaire evaluates the system's design focusing on key risk factors like fairness, transparency, and human oversight. The risk level assigned to each question is averaged into Specific Risk Indicators for each section, which are then averaged into the Questionnaire Risk Indicator (QRI), which ranges from 1 (low risk) to 10 (high risk). Some questions include a "blocking" answer, which, if selected, immediately halts the assessment. This response signals the need for system modifications, as it highlights design choices that could lead to unacceptable impacts. Table 2 provides an illustrative example of the type of questions addressed during the questionnaire phase of the FRIAct methodology. These questions are structured to ensure a thorough analysis of the AI system's potential impacts on fundamental rights, both before deployment and throughout its operational lifecycle. By addressing key technical, contextual, and legal considerations, this framework's purpose is to facilitate the implementation of a proactive approach to safeguarding fundamental rights in compliance with the AI Act. Thus, the Questionnaire provides essential information about the AI system's purpose, design, operational context, and affected populations, which serves as the foundation for Matrix analysis. For example, as it will be highlighted in Section 4.2, the technical details captured in Phase 2, such as algorithm type, data governance practices may have an impact on the calculation of Probability of Occurrence, while human oversight safeguards may impact the score assigned to Severity.

Then, the QRI summarises the AI system's risk profile across multiple dimensions (e.g., transparency, fairness, oversight). This QRI is directly used to individually weigh the significance of risks calculated through the Matrix, i.e. IS scores. While the Questionnaire contextualises the risks based on operational and technical factors, the Matrix provides numerical IS scores to detail the magnitude and likelihood of impacts on individual rights. Together, these tools offer a comprehensive framework that integrates qualitative insights with quantitative analysis, ensuring robust risk management and compliance under the AI Act. The following grid confronts some examples of the questions for Phase I and Phase II. While Phase I does not lead to obtaining a risk level, Phase II instead contains closed-ended questions with a risk score assigned to each possible answer.

Table 2 – Example of questions in the Questionnaire.

<i>Questions Phase I</i>	
<i>Categories</i>	<i>Questions</i>
General Context	Describe the context and the specific purpose of the AI system.
High-Risk Classification	Is the AI system classified as High-risk under article 6 of the AI Act and why?
Affected Groups	Are there any categories of natural persons that can be considered as more vulnerable groups? What are their specific risks?
<i>Questions Phase II</i>	
Technical Deployment	What type of algorithm(s) or model have been used? Please specify your choice and provide detailed examples. - a non-self-learning algorithm in which humans specify the rules the computer must comply to (choose a risk level from 1 to 3) - a System that is entirely or partially based on a self-learning algorithm, where the machine itself is finding patterns in the data (choose a risk level from 4 to 10)
Data and Fairness	Are all input data (from internal and external sources), including third-party training data and data added by the Deployer, governed by data governance and data quality processes?

	- Yes, all, in compliance with GDPR, AIA and IP requirements (risk level 1) - Only partially (e.g., in case a model has been pre-trained by an external provider and there are no warranties about data used). Please specify (choose a risk level between 2 and 10) - No (blocking)
Transparency	Are the components of the AI System and their outputs explainable, interpretable and/or verifiable? - Yes, all the components are designed to be explainable, interpretable and/or verifiable. Describe what techniques are employed (choose a risk level from 1 to 3) - Not all the components of the AI System are explainable, interpretable and/or verifiable. Describe what techniques are employed (choose a risk level from 4 to 10) - No (blocking)
Human Oversight	What is the degree of automated decision-making in the AI system? - The decision is taken by a human being and the AI System provides only an additional layer of information (choose a risk level between 1 and 2) - The decision is made by the application of AI System are only executed after human review or approval (choose a risk level between 3 and 4) - The decision relies on the AI System, but it is possible for a human to override the outcomes (choose a risk level between 5 and 10) - The decision-making process is completely reliant on the AI System, without possibility for overrides (blocking)
Monitoring and Updates	If unfair behavior emerges, will it be possible to intervene to correct it? - Yes, it will be possible to intervene (risk level 1) - No, it has been not possible to correct the unfair behavior (blocking)

3.2 The Matrix

The second element on which the assessment relies is the Matrix, designed to evaluate the potential impact of an AI system on each fundamental right enshrined in the CFREU. By quantifying risks through structured dimensions and scores, the Matrix offers a granular, data-driven analysis of how an AI system might affect individuals and groups. It is built to ensure consistency, reliability, and adaptability throughout the two phases: pre-deployment and post-deployment.

This tool provides a semi-quantitative analysis of potential impacts on fundamental rights as protected and listed by the CFREU. The Matrix produces Impact Significance (IS) scores, which quantify the risk to specific fundamental rights. To ensure reliable results, the data used in the Matrix must: a) extend beyond training datasets (e.g., validation and test datasets) to represent the system's behaviour more comprehensively and reduce overfitting risks; b) be representative of the real population affected by the deployed AI system, as required by Article 10 of the AI Act.

The Matrix is structured to consider up to all 50 rights enshrined in the EU Charter. Each right is evaluated separately to ensure a comprehensive assessment. The result is a distinct IS score for each right, representing the degree to which the AI system impacts that particular right. The compilation of the Matrix for FRIAct has been supplemented by a *vademecum* (see Appendix 2) proposed to enhance clarity and provide structured guidance for evaluating which fundamental rights might be impacted by an AI system. It aims to translate abstract principles of fundamental rights into actionable standards relevant to the AI context. Thus, the *vademecum* includes concise explanations for each right alongside targeted questions that probe the AI system's potential impact. These questions provide a systematic means to assess risks to individual rights. By integrating the *vademecum* into the FRIAct process, the matrix outputs – namely the IS scores – are rendered more precise and contextually relevant. Table 3 shows examples of Articles 1, 7, 8 and 47 as developed and clustered in the *vademecum*. The full version of the *vademecum* is reported in Appendix 2.

Table 3 – Excerpt from the *vademecum* for the Matrix compilation.

<i>Fundamental Right</i>	<i>Short Explanation</i>	<i>Example Guiding Questions</i>
Article 1: Human Dignity	Forms the basis of EU values; impacts privacy, equality, and social security.	Does the system respect individuals' autonomy? Does it avoid reinforcing harmful stereotypes or biases?

Article 7: Privacy	Ensures respect for private and family life; critical for gender equality and child protection.	Does the AI system collect personal data responsibly? Does it prevent unauthorized disclosure of private information?
Article 8: Data Protection	Protects data processing fairness and user access to rectification.	How does the system ensure data accuracy? Are processing practices transparent and aligned with GDPR requirements?
Article 47: Fair Trial	Guarantees effective remedies and fair judicial processes.	Does the system provide clear reasoning for its decisions? Are there redress mechanisms for contesting AI outputs?

The Matrix evaluates risks based on two key dimensions: severity,⁷⁹ hence the magnitude of the harm; the probability of occurrence,⁸⁰ as the likelihood of harm. To better simplify and to encapsulate the broadest picture possible on the potential infringement of fundamental rights, severity is composed of two sub-dimensions:

- **Intensity:** Measures the magnitude of potential harm, considering worst-case outcomes. It is the conceivable level of harm or damage that could result from a given risk. It considers the most severe possible outcome and classifies the gravity of that outcome based on its impact on a fundamental right.
- **Effort of Remediation:** The range of measures and resources required to address and potentially reverse any unintended or harmful outcome produced by the AI system.

Likewise, the probability of occurrence can be broken down in:

- **Likelihood:** The probability that the AI system will produce an error leading to an adverse impact on fundamental rights.
- **Robustness:** The specific performance indicator of the AI system. This sub-dimension is considered only in the post-deployment Matrix, as it requires post-deployment performance data.

Each sub-dimension is assigned a score on a 1 to 10 scale, in particular:

⁷⁹ For the purpose of our analysis, we define severity as “the extent or degree of harm on the protection of fundamental rights of a natural person or group that could occur if a risk materialize”.

⁸⁰ For the purpose of our analysis, we define the probability of occurrence dimension as “the likelihood to impact a fundamental right of a natural person or group of people due to the specific AI system performance”.

- **Intensity**

- 1–2 (Negligible)

Harm is minimal and easily mitigated.

- 3–4 (Moderate)

Noticeable but generally localized; potential impact remains reversible.

- 5–6 (Serious)

Significant effects on an individual's well-being or rights.

- 7–8 (Severe)

Major consequences, potentially involving long-term or broad-scale harm.

- 9–10 (Catastrophic)

Profound, possibly irreversible impact on a fundamental right.

- **Effort of Remediation**

- 1–2 (Trivial)

Corrective actions are minimal; existing resources suffice.

- 3–4 (Modest)

Remediation requires moderate effort and some coordination.

- 5–6 (Substantial)

Demands notable resource allocation, possibly specialized expertise.

- 7–8 (High)

Complex, time-intensive, and may disrupt operations.

- 9–10 (Nearly Impracticable)

Extremely difficult or expensive to address, risking feasibility issues.

- **Likelihood**

- 1–2 (Rare)

Very low probability; would occur only under exceptional conditions.

- 3–4 (Unlikely)
Possible but not expected in typical scenarios.
- 5–6 (Moderate)
Could happen regularly if certain factors align.
- 7–8 (Likely)
Reasonably anticipated to occur in normal operating conditions.
- 9–10 (Almost Certain)
Highly probable to occur unless major safeguards are implemented.

Robustness sub-dimension score is obtained using the complementary of the AI system's performance indicator. For example, if the system's accuracy is 80% (scored as 8 out of 10), then the Robustness sub-dimension score is calculated as $10 - 8 = 2$. As such, it is a truly quantitative measure. The scores for sub-dimensions are then combined linearly to calculate:

$$\text{Severity}(j) = \frac{\text{Intensity}(j) + \text{Effort of Remediation}(j)}{2}$$

$$\text{Probability of Occurrence}(j) = \frac{\text{Likelihood}(j) + \text{Robustness}(j)}{2}$$

Where (j) indicates that each dimension is referred to a specific fundamental right.

The IS score for each fundamental right is then obtained as:

$$\text{IS}(j) = \text{Severity}(j) \times \text{Probability of Occurrence}(j)$$

When assessing the risks to the individual rights listed in the CFREU, each right is assigned an IS score, which ranges from 1 to 100, with 1 indicating minimal risk and 100 representing maximum risk. Table 4 is a representation of the Matrix including the first 10 rights of the CFREU.

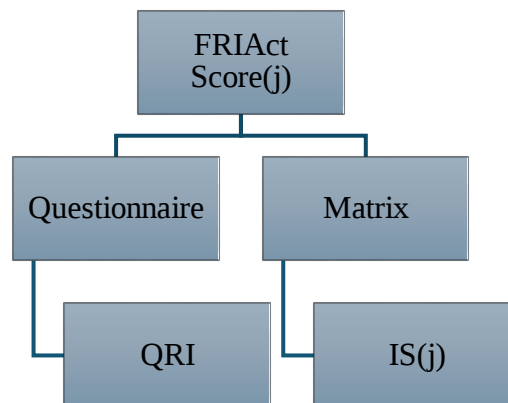
Table 4 – Matrix visual representation.

Fundamental Rights	Art.	Severity			Probability of occurrence (PO)			Impact Significance
		<i>Intensity</i>	<i>Effort of remediation</i>	Severity Level	<i>Likelihood</i>	<i>Robustness</i>	PO Level	
CHAPTER 1: DIGNITY								
Human dignity	1							
Right to life	2							
Right to the integrity of the person	3							
Prohibition of torture and inhuman or degrading treatment or punishment	4							
Prohibition of slavery and forced labour	5							
CHAPTER 2: FREEDOMS								
Right to liberty and security	6							
Respect for private and family life	7							
Protection of personal data	8							
Right to marry and right to found a family	9							
Freedom of thought, conscience and religion	10							

3.3 FRIAct: Integration of Questionnaire and Matrix

The Questionnaire and Matrix are designed to work together as core components of the FRIAct framework, ensuring a comprehensive evaluation of AI systems. While the Questionnaire captures qualitative insights and contextual understanding, the Matrix provides semi-quantitative, rights-specific analysis. The integration of the Questionnaire (QRI) and Matrix (IS) results in the FRIAct Score, which provide actionable insights into the AI system’s compliance and areas needing improvement. The third phase synthesizes the QRI scores with the outputs of the Matrix (IS) to calculate the FRIAct Scores for individual fundamental rights.

Figure 2 – Flowchart displaying the relationship between the Questionnaire and the Matrix within the FRIAct framework.



To produce a final, rights-specific, score, we need to combine QRI and IS(j) into a single index. We choose to do that by a weighted sum, where we decide to give more importance to IS(j), since this is the index capturing more granular, rights-specific, features.

The actual weights we are proposing below (30/70) are of course arbitrary, and can be intended as a starting point subject to further refinements and analysis. The same is true for the specific value of the thresholds that we propose, that trigger different actions in response to the assessment.

- **Calculation:** each FRIAct Score is derived using the following formula:

$$\text{FRIAct Score (FR(j))} = 0,3 \times \text{QRI(\%)} + 0,7 \times \text{IS(j)}$$

Where:

The QRI score is rescaled on a 1 to 100 scale in order to combine it with the IS(j) score; FR(j) refers to the FRIAct score for a specific fundamental right from the CFREU.

The IS(j) is the risk level associated with fundamental right (j) as calculated in the Matrix.

- **Thresholds and Outcomes:**

- FRIAct Score < 50: The system passes the assessment.
- FRIAct Score between 50 and 75: The system requires mitigation or monitoring actions before development can proceed.
- FRIAct Score $\geq$ 75: The system cannot be developed as planned and must be reevaluated.

- **Mitigation and Reporting:**

- Any section with a Risk Indicator of 75 or higher must be highlighted and described in detail before closing the assessment.
- For scores requiring mitigation (50–75), implement the necessary actions and rerun Phases 2 and 3 to verify compliance.

Thus, the integration of the Questionnaire and Matrix culminates in the calculation of FRIAct Scores for each fundamental right. These scores provide actionable insights for deployers and regulators by highlighting specific rights at risk and guiding mitigation strategies. The two instruments are intertwined in every phase. Hence, during the design phase, the Questionnaire identifies potential risks, and the Matrix evaluates projected impacts. Together, they guide system adjustments before deployment. Then, post-deployment, the Questionnaire updates contextual and operational insights, while the Matrix recalculates IS scores using real-world performance data.

The Questionnaire integrates the quantitative part of the FRIAct, gathering contextual and operational data about the AI system, including its intended purpose, affected populations, and technical characteristics. These insights are essential for setting the parameters used in the Matrix. The Questionnaire produces a QRI, summarising the system's overall risk profile. This QRI serves as a baseline for the Matrix by weighting its calculations of the IS scores. The QRI helps align qualitative assessments (e.g., operational context, technical safeguards) with the numerical risk levels calculated in the Matrix. Then, this latter one translates the general risk areas identified in the Questionnaire into precise, measurable impacts on individual fundamental rights.

- Example: if the Questionnaire identifies challenges in data governance, we expect that the Matrix will assign higher Likelihood scores under the Probability of Occurrence dimension, potentially increasing the IS score for rights such as privacy or data protection.

This iterative process ensures that the AI system adapts to evolving risks and maintains compliance throughout its lifecycle. This combination will be further heightened in the case scenario analysis that follows, demonstrating how the Questionnaire and Matrix, even though being interdependent tools that bridge qualitative and quantitative analyses, together ensure a thorough evaluation of AI systems under the FRIAct framework. By leveraging their combined outputs, deployers can proactively address risks to fundamental rights, comply with the AI Act, and foster ethical and responsible AI governance.

3.4 FRIAct Lifecycle

With respect to instructions related to the frequency of the assessment, it is advised to perform the FRIAct in both the two phases for every AI system subject to assessment: once during the pre-deployment phase and periodically thereafter through the post-deployment phase.

Table 5 – Confrontation between pre-deployment and post-deployment phase.

<i>Pre-deployment phase</i>	<i>Post-deployment phase</i>
It is recommended to initiate this phase as early as possible in the AI system's lifecycle, before deployment. Early-stage assessments facilitate the development of trustworthy systems by allowing adjustments while software development remains flexible. This can prevent potential infringements on fundamental rights and mitigate financial risks, ensuring the feasibility of deployment strategies.	After deployment, systems should undergo regular reassessments to account for changes in performance and operational contexts. Post-deployment monitoring allows organizations to verify assumptions made during the design phase and adapt to real-world conditions. Monitoring is not merely a verification of initial assumptions; it leverages a broader set of parameters derived from operational performance, providing a more detailed and realistic evaluation. It is suggested to conduct the post-deployment assessment at least every 12 to 18 months, with shorter intervals for cutting-edge technologies or systems that scored high in the FRIAct pre-deployment phase.

The FRIAct is flexible and can be tailored to fit the organisational structure and needs of the deploying entity. It provides a bridge between legal, ethical, technical, and domain-specific considerations, ensuring holistic compliance with the AI Act.

By following this structured, iterative approach, organisations can ensure that their AI systems not only comply with the AI Act but also align with ethical and societal values, safeguarding fundamental rights at every stage of the system's lifecycle. This approach provides a robust

mechanism for navigating the complexities of AI governance while fostering trust and accountability.

4. LoanLens: a testing ground use case

The following section illustrates how the FRIAct can be applied to a practical scenario. The use case involves an AI system used for credit scoring, a high-risk area under the AI Act,⁸¹ due to its potential impact on individuals' fundamental rights. This example, while fictitious, is designed to be plausible and representative of real-world applications.

4.1 Overview of the AI system

The AI system that we are considering has the purpose of calculating the credit scoring of natural persons. The prospected scenario consists of a Bank that currently relies on a traditional Machine-Learning System (MLS) to calculate a credit score for individual customers. This system draws on four primary indicators:

1. *A rating derived from 100 creditworthiness variables* - this composite rating is built from a wide array of data points reflecting an individual's financial behavior, such as historical repayment patterns, outstanding debt levels, credit utilization ratios, length of credit history, and other factors that statistically correlate with credit risk;
2. *An early warning system* - this mechanism is designed to continuously monitor customers' activities and external signals that might hint at a decline in creditworthiness. For example, if a pattern of missed payments emerges or there is adverse news about a customer's employer, the early warning system raises alerts. By flagging these signals promptly, the Bank can take proactive measures, such as requesting updated documentation or revising credit limits, to mitigate potential losses;
3. *Transactional or affordability data* - these data reflect the customer's ongoing financial obligations and day-to-day cash flows. Typical examples include monthly spending patterns, recurring bills, salary deposits, and other income streams. By assessing whether the borrower's incoming funds can support their current (and potentially additional) debt, the system can estimate affordability and likelihood of timely repayment;

⁸¹ Annex III, 5(b)

4. *A metric representing financial wealth* - this indicator provides insight into the customer's broader financial standing, often accounting for assets such as real estate, investment portfolios, or savings. A higher wealth metric generally suggests more capacity to handle debt, while a lower metric may signal limited resources to manage unforeseen financial strains.

Building on this existing infrastructure, the Bank proposes to integrate a conversational Decision Support System (DSS) powered by GenAI.

In practice, the credit score computed by the MLS is fed into the GenAI model along with various unstructured data sources (e.g., documents submitted during the credit application and historical customer profile data). Through a chatbot interface, called LoanLens, a human decision-maker gains a single, centralized access point for receiving credit score outputs and requesting additional analysis and information about customers applying for loans. The AI System's output will be a numerical score ranging from 1 to 100, where 1 is the lowest Credit Score and 100 is the highest achievable score.

To aid quick and consistent decision-making, the Bank categorizes this numerical score using a traffic light threshold framework:

- Red Zone (1–40): High credit risk, possibly requiring additional collateral or higher interest rates.
- Yellow Zone (41–70): Moderate credit risk, suggesting further verification steps or more stringent monitoring.
- Green Zone (71–100): Low credit risk, indicating eligibility for more favorable terms and expedited approval.

The person who makes the final decision on the loan application can rely on the given score or override it, providing a justification that explains the reasons for the disagreement. In this scenario, we assume that the GenAI component is a pre-trained model, developed by an external provider and then fed with purpose-specific data.

The GenAI model complements the MLS by synthesizing additional insights and providing a more user-friendly interface for human decision-makers. Here's how it works:

1. Data flow:

- The MLS generates a credit score based on its structured input indicators.
- This credit score, along with unstructured data (e.g., documents submitted during the credit application, historical customer profiles), is fed into the GenAI model.

2. Interface and functionality:

- LoanLens, the chatbot interface powered by the GenAI engine, acts as a single access point for human decision-makers.
- LoanLens allows managers to:
 - View the MLS-generated credit score.
 - Request deeper analyses or additional context about a customer's profile.
 - Interact in natural language to extract insights or locate relevant information efficiently.

3. Output:

- The AI System's output will be a numerical score ranging from 1 to 100, where 1 is the lowest Credit Score and 100 is the highest achievable score. To aid quick and consistent decision-making, the Bank categorizes this numerical score using a traffic light threshold framework:
 - Red Zone (1–40): High credit risk, possibly requiring additional collateral or higher interest rates.
 - Yellow Zone (41–70): Moderate credit risk, suggesting further verification steps or more stringent monitoring.
 - Green Zone (71–100): Low credit risk, indicating eligibility for more favorable terms and expedited approval.
- Human decision-makers can use this score as a suggestion but retain the authority to override it, providing a justification when doing so.

In this setup, the GenAI component is a pre-trained model provided by an external developer, further refined with purpose-specific data to ensure alignment with the bank's credit scoring needs.

The operational flow of the AI system ensures seamless integration between structured and unstructured data processing. In particular, the GenAI component:

- Utilises natural language prompts to process inputs and generate insights.
- Extracts relevant information from submitted documents, historical data, and credit scores.
- Synthesises this information into a user-friendly format, allowing managers to interact solely with LoanLens for all decision-support needs.

LoanLens simplifies access to layered information in one place, enabling quicker and more informed decision-making. In this sense, the natural language interface eliminates the need for technical expertise, making it accessible to a broader range of users.

A central design feature of this system is its emphasis on human oversight, as mandated by Article 14 of the AI Act. The following mechanisms ensure that human responsibility remains at the forefront:

1. Final decision:

- LoanLens provides suggestions and information, but all final decisions remain the responsibility of human decision-makers.
- Decision-makers can override or disregard the output, providing explanations for any discrepancies.

2. Stop mechanism:

- At any point, LoanLens can be stopped by the user to prevent unintended consequences, satisfying the Article 14(3) of the AI Act requirement for a 'stop button' feature.

3. Risk mitigation:

- By keeping the GenAI system in a supportive, rather than fully autonomous, role, the design minimises potential risks to fundamental rights, such as privacy violations or discrimination.

To ensure reliability and robustness, both the MLS and DSS components are tested rigorously before full deployment. These metrics are essential for compiling the post-deployment Matrix, particularly for assessing the Robustness sub-dimension:

1. MLS Performance:

- The MLS must achieve a classification accuracy of at least 80% in credit lending decisions during pre-deployment testing.
- This benchmark ensures the MLS reliably evaluates structured indicators like creditworthiness and financial health.

2. DSS Performance:

- The DSS must demonstrate at least 90% accuracy in responding to human-posed queries or locating relevant documents within the system.
- This ensures the DSS provides meaningful, accurate support to decision-makers.

3. Deployment Criteria:

- The system transitions to production only after both components meet their respective performance thresholds, ensuring a robust foundation for real-world operations.

4.2 The application of the FRIAct

In the following paragraphs we will show the most relevant results of the FRIA assessment conducted for LoanLens. The full Questionnaire is reported in Appendix 1.

In Phase 1, it is acknowledged LoanLens as a high-risk AI system under Article 6 and Annex III of the AI Act due to its role in determining access to credit, a private service with significant implications for fundamental rights. This classification is critical because credit scoring decisions can affect socioeconomic opportunities, such as housing and employment.

The system impacts two primary groups:

- Individuals applying for loans: These individuals directly rely on the AI-generated credit scores for access to credit services.

- Vulnerable groups: especially younger applicants that may have a ‘thin file’, i.e., having little or no credit history, and groups of individuals identified by sensitive characteristics potentially subject to algorithmic bias, such as foreigners, women, and persons with disabilities.

This phase establishes the broader context for LoanLens, emphasising its societal relevance and the need for stringent risk assessment.

Phase 2 dives into the technical components of LoanLens, assessing specific risks across key operational dimensions. Each dimension is assigned a Specific Risk Indicator, which contributes to the QRI. The full pre-deployment Questionnaire can be found in Appendix 1.

1. Deployment Process:

- LoanLens integrates a non-self-learning MLS and a self-learning GenAI component to process structured and unstructured data.
- The Risk Level for this dimension is 7.5, reflecting the complexity of combining these technologies and the systemic risks posed by the GenAI component.
- Alternative, simpler algorithms were considered but deemed insufficient due to lower accuracy, reinforcing the need for this advanced system.

2. Input Data and Fairness:

- The MLS complies with rigorous data governance standards, mitigating many risks. However, reliance on an external provider for the GenAI component introduces residual risks related to bias and transparency.
- This results in a Risk Level of 2.4, indicating moderate risk.

3. Transparency:

- The MLS outputs are explainable and interpretable, while the GenAI’s chatbot interface ensures verifiable insights for human decision-makers.
- This yields a Risk Level of 1.2, reflecting low risk in this area.

4. Human Oversight:

- The system maintains strong human oversight mechanisms, allowing decision-makers to override or halt the AI's operations, as required by Article 14 of the AI Act.
- A Risk Level of 1 indicates robust safeguards and minimal risk.

5. Monitoring and Maintenance:

- LoanLens includes monthly monitoring protocols and emergency update mechanisms to address emerging issues. These processes ensure continued compliance and adaptability.
- The Risk Level for this dimension is 2.5, representing moderate risk.

Table 6 – Pre-deployment Questionnaire Results

<i>Section</i>	<i>Risk Indicator (RI)</i>
Deployment Process	7.5
Input data and Fairness	2.4
Explainability and Transparency	1.2
Performance	4.0
Human Oversight	1.0
Monitoring and Maintenance	2.5
Questionnaire Risk Indicator (QRI)	3.1

The Matrix quantifies the system's impact on fundamental rights by evaluating Severity and Probability of Occurrence for each relevant right. Nine rights of the CFREU – Art. 1, Human Dignity, Art. 6 Right to liberty and security, Art. 9 Right to marry and right to found a family, Art. 16 Freedom to conduct a business, Art. 17 Right to property, Art. 25 The rights of the elderly, Art. 26 Integration of persons with disabilities, Art. 33 Family and professional life, Art. 36 Access to services of general economic interest – that have been evaluated by us as potentially impacted by the AI system in this context. Table 7 a visual of the Matrix, calculated on the nine mentioned fundamental rights.

Table 7 – LoanLens pre-deployment Matrix.

<i>Art.</i>	<i>Fundamental Right</i>	<i>Severity</i>			<i>Probability of Occurrence (PO)</i>		<i>IS (%)</i>
		<i>Intensity</i>	<i>Effort of Remediation</i>	<i>Severity Level</i>	<i>Likelihood</i>	<i>PO Level</i>	
1	Human Dignity	10	2	6	1	1	6
6	Right to liberty and security	6	2	4	1	1	4
9	Right to marry and right to found a family	6	2	4	1	1	4
16	Freedom to conduct a business	10	2	6	1	1	6
17	Right to property	10	2	6	1	1	6
25	The rights of the elderly	7	2	4.5	1	1	4.5
26	Integration of persons with disabilities	10	2	6	1	1	6
33	Family and professional life	8	2	5	1	1	5
36	Access to services of general economic interest	10	2	6	1	1	6

Despite high Severity scores for some rights, strong control measures and human oversight safeguards reduce the Probability of Occurrence, keeping overall IS values low.

Before addressing the completion of individual entries, it may be useful to investigate the values for Effort of Remediation and Likelihood. As it is possible to see, the values chosen are always the same: 2 for Effort of Remediation and 1 for Likelihood. These values are determined by the strong presence of the human oversight component. Therefore, we expect that both the effort required to

override an impactful decision and the frequency of such errors occurring will be significantly reduced due to the strong human oversight component.

Upon analyzing Article 1 CFREU, it is possible to notice that an intensity rating of 10/10 has been assigned. This rating reflects our assessment that a wrongful decision to deny a loan would constitute a severe violation of an individual's fundamental right to personal dignity, which we deem to be of the utmost importance. This determination is based on the ethical principle that such decisions, if erroneous, can significantly undermine an individual's sense of agency and personal worth, thus warranting the maximum possible intensity for this particular risk.

Moving to the analysis of Article 6 'right to liberty and security', an intensity of 6/10 has been chosen. This choice is justified by the fact that an erroneous denial of credit could undermine both the security and the freedom of individuals. A slightly above-average intensity was selected because the economic capacity of individuals enables access to essential services that are important for the full enjoyment of personal security and freedom rights.

We followed the same considerations while assessing the impact of the AI systems with regards to Article 9 'right to marry and found a family': we have again chosen an intensity of 6 out of 10. We considered that, for the full enjoyment of the aforementioned Article 9, a proper allocation of loans is necessary. In the case of personal loans, we deemed this article potentially impacted, as it protects the right to be free (and thus 'able') to find a family.

On the other hand, given its close connection to the economic sphere of the individual, we have selected an intensity of 10 out of 10 for the potential impact on the freedom to conduct business (Article 16). Since this freedom is directly linked to the financial situation of the loan applicant, we expect that an erroneous denial of the loan would have the maximum possible impact on the individual's circumstances. The same considerations apply also to article 17 'right to Property'.

The aim of Article 25 is to combat social exclusion and discrimination against the elderly, fostering a society where older individuals are valued as a resource and can live actively and with respect. Based on these considerations, we have assigned an intensity value of 7 out of 10. This decision reflects the potential limitations in exercising this right due to wrongful decisions made by AI systems. Additionally, economic factors may hinder the elderly's ability to participate in activities of interest, thereby undermining the protections outlined in Article 25 of the Charter.

Article 26 CFREU establishes the right to integration for persons with disabilities. Based on considerations related to the socio-economic status of this particularly vulnerable group, an intensity rating of 10 out of 10 has been selected once again. Given the additional costs these individuals must bear to effectively exercise their rights, we anticipate that an unjustified denial of a loan could have the maximum possible impact on this group.

The same considerations made for, among others, Article 16 have been applied to Article 36, which defines the right to access services considered to be of general economic interest. This is due to the nature of the credit requested and the connection between the right enunciated in Article 36 and the economic sphere outlined by this right.

Regarding Article 33 of the Charter, ‘family and professional life’, an intensity value of 8 out of 10 was assigned. This is due to the following considerations: being wrongly denied a personal loan can lead to the consequence that the balance between work and family life, as protected by Article 33, is significantly undermined. Financial stability results as a core factor in the fulfillment of Article 33.

To conclude the considerations regarding the choices made during the completion of the pre-deployment Matrix, it is important to focus on the IS of the individual articles. As can be seen, the final results are all low scores (ranging from an IS of 4% to an IS of 6%) despite the significantly high intensity values. This is possible due to the design decisions that were made, in addition to the strong human oversight involved in determining the final output. This control is appropriately positioned between the production of the output and the individual potentially impacted by it. Such intermediation allows for quicker detection of errors in the assessment of the loan applicant and ensures the prompt intervention of the designated manager responsible for this task.

Table 8 – Pre-deployment FRIAct Score.

<i>Art.</i>	<i>QRI (%)</i>	<i>IS (%)</i>	<i>FRIAct Score (%)</i>
1	31	6	13.5
6	31	4	12.1
9	31	4	12.1
16	31	6	13.5
17	31	6	13.5
25	31	4.5	12.45

26	31	6	13.5
33	31	5	12.8
36	31	6	13.5

Post-deployment, the Questionnaire incorporates real-world implementation informations. The full post-deployment Questionnaire can be found in Appendix 1. Results are shown in the following table:

Table 9 – Post-deployment Questionnaire results.

<i>Section</i>	<i>Risk Indicator (RI)</i>
Deployment Process	7.5
Input data and Fairness	2.4
Explainability and Transparency	1.2
Performance	4.0
Human Oversight	1.0
Monitoring and Maintenance	2.5
Ownership and Control	10
Questionnaire Risk Indicator (QRI)	4.1

The post-deployment Matrix takes into account also Robustness sub-dimension. Please note that Intensity, Effort of Remediation and Likelihood sub-dimensions are expressed as numerical indicators ranging from 1 to 10, where 1 represents the lowest level and 10 represents the highest level of Intensity, Effort, or Likelihood. By contrast, the Robustness sub-dimension is measured as the inverse of the AI system's performance—thus, to maintain consistency with the 1–10 scale, a score of 1 indicates the highest level of Robustness, and a score of 10 indicates the lowest level of Robustness. For this toy example, as described above, the ML component has to achieve an accuracy of 80% to be effectively deployed, so Robustness indicator is:

$$\text{Robustness} = 10 - \text{Accuracy} = 10 - 8 = 2$$

This adjustment slightly increased IS scores compared to the pre-deployment Matrix, reflecting greater precision in post-deployment evaluations.

Table 10 – Post-deployment Matrix results.

Art.	Severity			Probability of Occurrence (PO)			IS (%)
	<i>I</i>	<i>ER</i>	<i>S</i>	<i>L</i>	<i>R</i>	<i>PO</i>	
1	10	2	6	1	2	1.5	9
6	6	2	4	1	2	1.5	6
9	6	2	4	1	2	1.5	6
16	10	2	6	1	2	1.5	9
17	10	2	6	1	2	1.5	9
25	7	2	4.5	1	2	1.5	6.75
26	10	2	6	1	2	1.5	9
33	8	2	5	1	2	1.5	6.75
36	10	2	6	1	2	1.5	9

Compared to pre-deployment Matrix results, IS indicators resulting from post-deployment Matrix are slightly higher than the ones obtained from the pre-deployment one. Taking into account the actual performance of the AI system led to a very slight worsening of the scores, which were nevertheless extremely low due to the centrality of the designed human oversight component. Post-deployment FRIAct Scores are described in the following table.

Table 11 – Post-deployment FRIAct Scores.

Art.	QRI (%)	IS (%)	FRIAct Score (%)
1	41	9	18.6
6	41	6	16.5
9	41	6	16.5
16	41	9	18.6
17	41	9	18.6
25	41	6.75	17.03

26	41	9	18.6
33	41	6.75	17.03
36	41	9	18.6

Lifecycle adjustments are essential for maintaining compliance. The system is reassessed in cases of significant modifications, such as changes to the GenAI model or new deployment contexts. These iterative assessments ensure LoanLens evolves responsibly and aligns with the FRIAct framework's standards.

The LoanLens use case highlights the FRIAct framework's ability to integrate the qualitative insights of the Questionnaire with the quantitative approach of the Matrix. By systematically evaluating risks across operational and rights-based dimensions, the framework identifies potential vulnerabilities and guides mitigation strategies. Robust human oversight mechanisms, coupled with continuous monitoring, ensure that the system complies with the AI Act throughout its lifecycle. This approach not only safeguards fundamental rights but also fosters trust and accountability in AI deployment.

5. Conclusion

This paper presents the FRIAct framework as a comprehensive approach for assessing the impact of AI systems on fundamental rights, particularly in compliance with the AI Act. The framework is rooted in the principles enshrined in the CFREU and operationalised through a combination of qualitative and quantitative tools, specifically the Questionnaire and the Matrix. Together, these instruments provide a rigorous structure for evaluating the risks posed by AI systems, particularly high-risk applications such as those involving GenAI and other advanced technologies.

The FRIAct framework is more than a procedural tool. It represents a systematic effort to align AI innovation with European constitutional values. By embedding the assessment of fundamental rights into the lifecycle of AI systems, the framework bridges the gap between the theoretical commitments to rights protection and the practical realities of AI deployment. This approach is crucial in addressing the transformative and sometimes disruptive effects of AI on individuals and society.

The Questionnaire is designed to provide a qualitative understanding of an AI system's context, purpose, and deployment, with a particular focus on identifying risks to fundamental rights. It

establishes the framework for evaluating its operational context and the populations it may impact. This qualitative groundwork is complemented by the Matrix, which adds a assessment specifically designed to produce a quantitative output by systematically mapping potential qualitative impacts of the system on specific rights, such as privacy, non-discrimination, and human dignity. The Matrix's numerical outputs, expressed as IS scores, are grounded in measurable dimensions of risk, including severity and probability of occurrence. This integration of qualitative and quantitative approaches ensures that the framework is both comprehensive and actionable.

The framework's flexibility makes it applicable to a wide range of AI systems, from traditional machine learning models to sophisticated GenAI systems. This adaptability is particularly evident in the LoanLens use case, which demonstrates how the FRIAct framework can be applied to assess the risks associated with a hybrid credit-scoring system that combines structured data analysis with generative AI capabilities. The case study illustrates how the framework addresses challenges such as transparency, human oversight, and fairness, while also emphasising the importance of ongoing monitoring to adapt to changes in system performance and regulatory environments. By incorporating robust safeguards and emphasising the need for human accountability, the framework aligns with the principles of Article 14 of the AI Act, which mandates human oversight of high-risk AI systems.

Another significant contribution of FRIAct is its ability to enhance compliance with the AI Act while fostering public trust in AI systems. The framework's lifecycle approach, encompassing both pre-deployment and post-deployment phases, ensures that risks are not only identified and mitigated during the development of AI systems but also continuously evaluated and managed throughout their operational lifespan. This dynamic risk management process is essential in a rapidly evolving technological landscape, where new risks may emerge as AI systems interact with diverse real-world contexts. The framework's emphasis on inclusivity further ensures that the rights of vulnerable populations, such as the elderly or individuals with disabilities, are carefully considered and protected.

FRIAct also addresses critical gaps in the current regulatory landscape by offering a practical pathway to implement the FRIA obligations under Article 27 of the AI Act. The framework's integration of the Questionnaire and Matrix not only aims to provide a clear approach to risk assessment but also sets a standard for aligning AI practices with broader societal values. This alignment is particularly important in high-stakes applications like credit scoring, where decisions

have far-reaching consequences for individuals' access to essential services and economic opportunities.

In addition to its practical utility, FRIAct underscores the importance of fostering a collaborative governance model. The framework highlights the need for greater cooperation between AI providers and deployers, addressing the information asymmetry that often hampers effective risk management. By ensuring that deployers have access to technical and operational insights provided by AI developers, FRIAct facilitates a more transparent and accountable AI ecosystem. Moreover, the framework emphasises the role of multidisciplinary teams, comprising legal, technical, and ethical experts, in implementing comprehensive assessments. This collaborative approach not only enhances the robustness of the assessment process but also strengthens the organizational capacity to navigate the complex regulatory requirements of the AI Act.

Looking ahead, the FRIAct framework offers a pathway for policymakers, regulators, and organisations to operationalise the principles of ethical and responsible AI. As the EU continues to refine its AI governance structures, frameworks like FRIAct can serve as benchmarks for implementing the AI Act's provisions while ensuring that AI systems respect and uphold fundamental rights. The framework's adaptability, rigorous methodology, and emphasis on lifecycle risk management position it as a critical tool for fostering trust and accountability in AI systems.

In conclusion, the FRIAct framework exemplifies how AI governance can balance innovation with the protection of fundamental rights. By providing a structured, replicable, and context-sensitive approach to risk assessment, the framework not only meets the compliance needs of the AI Act but also contributes to building a rights-centred digital ecosystem.

Acknowledgements

The authors would like to thank Luca Paulicelli (Intesa Sanpaolo) for his precious contribution in designing the LoanLens case study, Valerio Cencig and Mario D'Almo (Intesa Sanpaolo) for insightful discussions on a number of topics related to the manuscript, Marco Ditta and Massimo Proverbio (Intesa Sanpaolo) for their constant help and support on themes of Responsible AI.

This work has partially been supported by the PNRR-M4C2 project "FAIR - Future Artificial Intelligence Research" funded by the European Union.

APPENDIX 1 – Questionnaire for LoanLens

This appendix contains the pre-deployment and post-deployment Questionnaires filled for the LoanLens case study (section 4).

Pre-deployment Questionnaire

Phase 1 - Context and purpose

A. General information:

- a. Describe the context and the specific purpose of the AI system.*

The Bank plans to enhance its existing infrastructure (traditional MLS) by integrating a GenAI-powered conversational DSS. LoanLens will combine the computed credit score with unstructured data (e.g., documents collected during loan processing and historical customer data with 10-year historical depth). LoanLens will be a chatbot interface, with which humans will be able to dialogue in a simple and fast manner. The specific purpose of the AI system is then to generate a credit score of natural persons to be used as “suggestions” for loan decisions, and to facilitate data extraction and information synthesis about customers, through a friendly interaction in natural language between humans and LoanLens.

- b. When (dd/mm/yyyy) will the AI system begin operation?*

The deployment of the AI system will occur on 03/06/2025.

- c. Specify the end date if known or estimate the duration in months/years.*

We estimate that the AI system will be used for at least three years.

- d. Usage Frequency: How many times per day/week/month will the AI system be used? Provide an exact count.*

The AI system will be used for all the loan applications that the Bank receives from retail customers (natural persons). It is not possible to give an exact number, but based on the history, it is expected that around one thousand requests per week will be processed.

- e. *What are the expected outputs of the AI system and how will they be used in decision-making processes?*

The expected output will be a numerical score ranging from 1 to 100, representing the creditworthiness of customers. This credit score will be available through LoanLens interface, together with the possibility of asking for additional information and pieces of supporting documents. The final decision remains human, and operators can decide autonomously whether to 'follow' the AI system's suggestion, or to deviate by providing an explanation.

- f. *Is the AI system classified as High-risk under article 6 of the AI Act and why?*

Yes, the LoanLens has been classified as a high-risk AI system since it is used to produce a credit score of natural persons and therefore determining access to and enjoyment of essential private services as stated in Annex III, 5 (b) of the AI Act.

- g. *Does the AI system have a GenAI component?*

Yes, LoanLens includes a GenAI component, which allows the answers to be generated. We specify that the credit score remains calculated by the traditional MLS and only 'reported' in natural language by the chatbot. .

- h. *Does the GenAI component present systemic risks as defined by the AI Act, and if so, what are the specific factors contributing to this assessment?*

No, it does not fall within Art. 3, par. 65 of the AI Act.

B. Categories of natural persons and groups

- a. *Who are the affected persons?*

The affected persons are all loan applicants, belonging to the 'retail' segment of customers (natural persons).

- b. *Are there any categories of natural persons that can be considered as more vulnerable groups? What are their specific risks?*

The disproportionate vulnerability between different subgroups of people in AI credit scoring application is a phenomenon widely analysed in the literature, with many points of attention raised. AI-driven credit scoring poses significant risks of exacerbating existing inequalities and creating new ones, particularly for the following subgroups of people:

- Youngest People. Young individuals, especially those just entering the workforce or with limited credit histories, often possess "thin files." AI systems, reliant on

extensive data, may struggle to accurately assess their creditworthiness, leading to unfair denials or inflated interest rates.

- **Elderly People.** Elderly individuals may face unique challenges, such as declining income due to retirement, increased healthcare expenses, or denials of opportunities due to their life expectancy. AI systems, may not adequately account for these life-stage shifts, resulting in an unfair risk assessment. Also, older people may not have as much digital credit-related footprint as younger generations, that could result in less data for the AI to work with.
- **People with Disabilities.** Individuals with disabilities may experience employment instability, fluctuating income, increased healthcare costs, or denials of opportunities due to their life expectancy. AI systems, lacking nuanced understanding of these factors, could misinterpret these patterns as indicators of higher credit risk, leading to discriminatory outcomes. Moreover, AI systems could also factor in health data, if it is available, in a way that could be discriminating. The instability of employment, and health conditions, can cause patterns in financial history that are not a sign of poor credit worthiness, but a sign of societal inequality.
- **Subgroups subject to biases due to sociodemographic characteristics.** AI systems are susceptible to inheriting and amplifying biases present in training data, which can disproportionately affect subgroups based on race, ethnicity, gender, or socioeconomic status. Therefore, historical data may reflect discriminatory lending practices, leading to biased AI-driven credit scoring. It is critical to be aware that historical data often contains societal bias, and AI systems may perpetrate or exacerbate those biases in a systematic way.

The groups outlined above are potentially more vulnerable to adverse outcomes in AI-driven credit scoring applications such as LoanLens, due to factors such as data scarcity, life-stage transitions, disability-related challenges, and the multi-faceted risk of algorithmic bias.

c. *Who will directly interact with the AI system?*

The AI system will only interact with the Bank's operators, credit specialists, through a conversational mode. There will be no direct interaction with the Bank's customers.

Phase 2.1 - Deployment process section

- a. What type of algorithm(s) or model will be used? Please provide detailed examples and context for each risk level.
- i. a non-self-learning algorithm in which humans specify the rules the computer must comply to (risk level 1 to 3)
 - ii. a System that is entirely or partially based on a self-learning algorithm, where the machine itself is finding patterns in the data (choose a risk level from 4 to 10)

RISK LEVEL: 8

COMMENT: The risk level has been chosen considering that the system is composed of a Gen AI component and a narrow AI component. The risk level chosen is 8 since the Gen AI components find patterns in the data.

- b. Among different alternatives suitable to achieve similar goals and performances, do you plan to use a simpler and more explainable algorithm?
 - i. Yes (choose a risk level between 1 and 5)
 - ii. No. Please explain your choice (choose a risk level between 6 and 10)

RISK LEVEL: 7

COMMENT: In this use case, we opted against using a simpler, more explainable algorithm due to performance considerations. While simpler methods may offer greater transparency, they typically cannot match the higher accuracy and efficiency provided by the more complex models.

Deployment process section risk level: 7.5

Phase 2.2 - Input data and Fairness section

- a. Are all input data (from internal and external sources), including third-party training data and data added by the Deployer, governed by data governance and data quality processes?
 - i. Yes, all, in compliance with GDPR, AIA and IP requirements (risk level 1)
 - ii. Only partially (e.g., in case a model has been pre-trained by an external provider and there are no warranties about data used). Please specify (choose a risk level between 2 and 10)
 - iii. No⁸²

RISK LEVEL: 4

COMMENT: The Narrow AI component is fully compliant with the relevant regulations and IP requirements, since it is designed and developed

⁸² If the Team selected this answer, the Questionnaire shall be suspended.

internally, so, minimizing risks. The GenAI component, being developed externally, introduces some risks related to compliance (which are mitigated by legal clauses) and intellectual property (which cannot be fully guaranteed). Thus, we have assigned a "risk level 4" based on the combination of these factors.

- b. If the AI system includes a GenAI component and it uses personal data to produce the output, could this lead to the generation of unexpected and/or undesired content? Do not count this question for the Risk Indicator calculation in case the System has no GenAI components or if it does not use personal data to produce the output.
 - i. Yes (risk level 10)
 - ii. No (risk level 1)

RISK LEVEL: 1

COMMENT: The GenAI component is designed to respond strictly based on the documents it has received as input. Consequently, it does not generate new or unexpected content; rather, it accelerates the user's access to relevant information without producing any automated decisions or recommendations.

- c. Have you or your Provider analyzed the input data to assess possible biases that could lead to a negative impact on fundamental rights? Please specify your choice and provide detailed examples.
 - i. Yes, please specify (e.g. in Fairness Assessment or data exploration, documentation provided by the Provider) (choose a risk level between 1 and 10).
 - ii. No⁸³

RISK LEVEL: 3

COMMENT: We conducted a direct assessment of the ML component's input data to identify and address any potential biases that could negatively affect fundamental rights. Additionally, the provider of the gen AI component has assured us that they perform a similar bias analysis on their input data, but we do not have direct access to them.

⁸³ If the Team selected this answer, the Questionnaire shall be suspended.

- d. Have you or your Provider collected all the relevant and available data needed to assess if the algorithm may discriminate against or disadvantage specific population subgroups (e.g., nationality, gender, age, etc.)? Please specify your choice and provide detailed examples.

i. Yes, please specify (e.g. in Fairness Assessment, data exploration, or in order to do this assessment; choose a risk level between 1 and 10)

ii. No⁸⁴

RISK LEVEL: 3

COMMENT: Regarding the ML component, we collected all the relevant and available data to assess if the algorithm may discriminate against or disadvantage specific population subgroups. Additionally, the provider of the Gen AI component has assured us that they perform a similar analysis on their input data, but we do not have direct access to them.

- e. Do you or your provider plan to monitor the trend over time of fairness metrics and/or AI System performance referring to specific population subgroups? Please specify your choice and provide detailed examples.

i. Yes (choose a risk level between 1 and 5)

ii. No (choose a risk level between 6 and 10)

RISK LEVEL: 1

COMMENT: We plan a monthly monitoring.

Input data and Fairness section risk level: 2.4

Phase 2.3 - Transparency section

⁸⁴ If the Team selected this answer, the Questionnaire shall be suspended.

- a. Are the components of the AI System and their outputs explainable, interpretable and/or verifiable ?
- i. Yes, all the components are designed to be explainable, interpretable and/or verifiable. Describe what techniques are employed (Choose a risk level from 1 to 3)
 - ii. Not all the components of the AI System are explainable, interpretable and/or verifiable. Describe what techniques are employed (Choose a risk level from 4 to 10)
 - iii. No⁸⁵

RISK LEVEL: 3

COMMENT: The ML component is explainable and interpretable. We chose risk level 3 due to the black-box nature of the Gen AI component, even though its outputs are verifiable.

- c. Have you identified the subjects (Provider, Deployer, and Affected persons) for which the output of the AI System shall be made sufficiently understandable?
- i. Yes (risk level 1)
 - ii. No⁸⁶

RISK LEVEL: 1

- d. Are the AI System outputs designed as sufficiently understandable for the Deployer?
- i. Yes (choose a risk between 1 and 10)
 - ii. No⁸⁷

RISK LEVEL: 1

- e. If applicable, are the AI System outputs designed as sufficiently understandable for the staff and other persons dealing with the operation and use of AI system on behalf of the Deployer? If not applicable, do not consider this question within the Risk indicator calculation.
- i. Yes (choose a risk between 1 and 10)

⁸⁵ If the Team selected this answer, the Questionnaire shall be suspended

⁸⁶ If the Team selected this answer, the Questionnaire shall be suspended

⁸⁷ If the Team selected this answer, the Questionnaire shall be suspended

- ii. No⁸⁸

RISK LEVEL: 1

- f. If applicable, are the AI System outputs designed as sufficiently understandable for the affected person? If not applicable, do not consider this question within the Risk indicator calculation.
 - i. Yes (choose a risk level between 1 and 10)
 - ii. No⁸⁹

RISK LEVEL: n/a

Transparency section risk level: 1.2

Phase 2.4 - Performance section

- a. Have you planned to test the AI system performance and proper functioning?

- i. Yes (choose a risk level between 1 and 10)
- ii. No⁹⁰

If yes, specify what is the metric chosen for that purpose and what is the frequency of the testing. On the basis of which evaluation criteria is a metric value considered acceptable (ex. acceptable thresholds)?

RISK LEVEL: 4

COMMENT: The ML component performance is assessed using accuracy metric as a reference. Gen AI component performance is assessed looking at the rate of correct

⁸⁸ If the Team selected this answer, the Questionnaire shall be suspended

⁸⁹ If the Team selected this answer, the Questionnaire shall be suspended

⁹⁰ If the Team selected this answer, the Questionnaire shall be suspended

responses or location of relevant documents at manager-posed queries. We plan to test the proper functioning on a monthly basis.

Performance section risk level: 4

Phase 2.5 - Human Oversight section

a. What is the degree of automated decision-making in the AI system?

- i. The decision is taken by a human being and the AI System provides only an additional layer of information (Choose a risk level between 1 and 2)
- ii. The decision is made by the application of AI System are only executed after human review or approval (Choose a risk level between 3 and 4)
- iii. The decision relies on the AI System, but it is possible for a human to override the outcomes (Choose a risk level between 5 and 10)
- iv. The decision-making process is completely reliant on the AI System, without possibility for overrides⁹¹

RISK LEVEL: 1

b. Can the operation of the AI system be interrupted through a 'stop' button or similar procedure?

- i. Yes, intervention possible (choose a risk level between 1 and 10)
- ii. No, intervention is not possible. Provide details⁹²

RISK LEVEL: 1

COMMENT: The human can interrupt the AI system in each moment of its interaction.

⁹¹ If the Team selected this answer, the Questionnaire shall be suspended

⁹² If the Team selected this answer, the Questionnaire shall be suspended

- c. Is it ensured that the staff and other persons dealing with the operation and use of AI systems on behalf of the Deployer and the Provider have the means to interact with the AI system and use it in an informed and conscious manner?

- i. Yes, those subjects have received a specific training and they can consult technical guidelines or support. Please specify. (choose a risk level from 1 to 4)
- ii. Yes, those subjects have received a specific training but they can not consult technical guidelines or support. Please specify (choose a risk level from 5 to 10)
- iii. No, those subjects have not been trained for the specific purpose⁹³

RISK LEVEL: 1

COMMENT: We provide specialized training for managers working with the AI system, offering intensive and mandatory sessions before they can use it. Additionally, we have a detailed set of guidelines for handling AI systems and offer technical support to managers in case of any doubts or malfunctions.

- d. Are there support mechanisms for the subjects mentioned above designed to address issues or concerns with AI system operation?
- i. Yes (risk level 1)
 - ii. No (risk level 10)

RISK LEVEL: 1

Human Oversight section risk level: 1

Phase 2.6 - AI system Monitoring and Maintenance

- a. If unfair behavior emerges, will it be possible to intervene to correct it?

⁹³ If the Team selected this answer, the Questionnaire shall be suspended

- i. Yes, it will be possible to intervene (risk level 1)
- ii. No, it has been not possible to correct the unfair behavior⁹⁴

RISK LEVEL: 1

- b. Do you plan to adopt monitoring techniques and related countermeasures in order to intercept anomalous behavior or performance deterioration?

- i. Yes (risk level between 1 and 10)
- ii. No⁹⁵

RISK LEVEL: 4

COMMENT: We plan to implement monitoring techniques across the entire AI system to detect any anomalous behavior or performance drops. While we can directly apply countermeasures to the ML component, we lack direct access to the Generative AI model provided by an external party. To address this limitation, we will employ a guardrail technique to ensure controlled and secure outputs.

- c. How do you plan to update the AI System (retrain, tune, knowledge source update, etc)?
 - i. Update not needed given the AI system's technical features (rule-based system without variations over time) (risk level 1)
 - ii. Not applicable, as in the case of GenAI Systems, either the AI system has been integrated by external providers (i.e. "proprietary models") or its update is economically too burdensome and operationally unnecessary (risk level between 2 and 9).
 - iii. Manual or automatic update planned (e.g., using new data, re-estimation of the hyperparameters of the model, vector store update in RAG applications.) (choose a risk level between 1 and 9)
 - iv. No update planned (risk level 10)

RISK LEVEL: 4

⁹⁴ If the Team selected this answer, the Questionnaire shall be suspended

⁹⁵ If the Team selected this answer, the Questionnaire shall be suspended

COMMENT: We chose risk level 4 because the gen AI component is periodically updated by the external party and not directly by us.

d. Have you planned procedures for emergency/urgent updates?

- i. Yes, please describe (choose a risk level between 1 and 5)
- ii. No (choose a risk level between 6 and 10)

RISK LEVEL: 4

COMMENT: We chose risk level 4 because we can immediately intervene on the ML component in case of an emergency, while we can't immediately intervene on the gen AI component.

e. Have you planned any processes in order to ensure the ongoing and continued availability of data, coherently to the business needs?

- i. Not needed since not foreseen by business requirements (risk level 1)
- ii. Yes (choose a risk level between 1 and 5)
- iii. No (choose a risk level between 6 and 10)

RISK LEVEL: 1

f. Do you plan to oversee the stability of input data/features used by the AI system?

- i. Yes, overseeing the stability of input data and features has been planned with one or more of these options (choose a risk level between 1 and 5):
 - 1. General and specific quality controls of input data
 - 2. Trend analysis of input data
 - 3. Trend analysis of Model features
 - 4. Specific Data Drift synthetic measures
 - 5. Evaluation of suitability/toxicity of prompts
- ii. Input data quality controls already in place are deemed as effective safeguards and no other actions are planned (choose a risk level between 6 and 9)

- iii. No action planned (risk level 10)

RISK LEVEL: 1

AI System Monitoring and Maintenance section risk level: 2.5

QRI: 3.1

Post-deployment Questionnaire

In this section we describe the post-deployment Questionnaire results. In order to avoid repetition, we will only present the additional question to those already present in the pre-deployment Questionnaire. Please note that in a real scenario the post-deployment Questionnaire is completed only after the AI system has been deployed, incorporating any new information that arises post-deployment. This post-deployment Questionnaire serves a different purpose than the one used during the design phase, so it must be filled out from scratch rather than relying on responses previously provided during the design stage.

Phase 2.7 - Ownership and Control

- a. Does the AI System use software libraries, and/or modules that have been developed by third parties (e.g., pretrained models, outsourced code, ...)? Please specify your choice and provide detailed examples.
 - i. No, AI System has been developed internally and/or using documented and accessible open-source libraries (e.g., pandas, scikit-learn) (risk level 1)
 - ii. Yes (risk level 10)

RISK LEVEL: 10

COMMENT: The AI system includes a gen AI component pre-trained and developed by an external party.

Ownership and control section risk level: 10

QRI: 4.1

APPENDIX 2 – Questions proposal for compiling the Matrix (*vademecum*)

This appendix contains a set of questions intended to support the compilation of the Matrix (section 3.2).

I. Methodology

To elaborate the framework of guiding questions to perform the FRIA, the analysis started to identify the core concepts beneath the single rights. The main sources to develop the short explanation and the guiding questions are the CFREU itself, as well as Commentaries to the Charter and the interpretation suggested by the jurisprudence of the Court of Justice of the European Union. What follows is the result of this background analysis.

II. Proposed questions

i. Art. 1 CFREU – Human dignity

- Short explanation:

Human dignity, as enshrined in Article 1 of the EU Charter of Fundamental Rights, forms the basis of fundamental rights and is integral to the European Union's values, including freedom, equality, and solidarity. It influences various other rights, impacting everything from privacy and equality to workers' rights and social security.

- Guiding questions:

1. **Autonomy and Decision-Making:** Does the AI system enhance or undermine the ability of individuals to make autonomous decisions? How does it affect individuals' control over their life choices?
2. **Dehumanization and Stereotyping:** Does the AI system risk treating individuals as mere data points, ignoring their unique human characteristics and needs? Does it reinforce negative stereotypes or biases that could degrade human dignity?
3. **Vulnerable Populations:** How does the AI system impact vulnerable groups such as the elderly, disabled, or minorities? Does it support their participation in social and cultural life or does it create barriers that might isolate or discriminate against them?

ii. Art. 2 CFREU – Right to life

- Short explanation:

The right to life, articulated in Article 2 of the EU Charter of Fundamental Rights, involves complex legal and ethical dimensions, particularly as they relate to new technologies and medical practices. The right not only prohibits arbitrary deprivation of life but also influences EU and MS laws and policies concerning healthcare access, emergency medical interventions, and ethical issues like euthanasia and assisted reproductive technologies.

- Guiding questions:

1. **Access to Healthcare:** How does the AI system manage or influence access to essential healthcare services and life-sustaining treatments? Does AI influence healthcare decisions in ways that could impact life-preserving treatments? Does the AI system enhance or impede access to urgent medical interventions that could be life-saving? Does the AI system enhance or impede access to specific medical treatment, like ATP?
2. **End-of-Life Decisions:** Does the AI system play any role in decisions around euthanasia or assisted dying, and how does it ensure respect for personal choices and dignity at the end of life?

iii. Art. 3 CFREU – Right to integrity of the person

- Short explanation:

Article 3 of the EU Charter of Fundamental Rights focuses on the right to physical and mental integrity, specifically emphasizing protections in the fields of medicine and biology.

- Guiding questions:

1. **Consent and Autonomy:** How does the AI system ensure that it upholds the patient's right to consent, especially in medical or biological contexts?

2. **Impact on Treatment:** Does the AI system involve procedures that could compromise the physical or mental integrity of individuals? Does the AI system influence medical decisions in ways that could compromise the physical or mental integrity of individuals?

iv. **Art. 4 CFREU – Prohibition of torture and inhuman or degrading treatment or punishment**

- **Short explanation:**

Article 4 of the EU Charter of Fundamental Rights prohibits torture and inhuman or degrading treatment or punishment. This is particularly relevant in the fields of justice, home affairs, and external policies where the EU and MS's influence is significant.

- **Guiding questions:**

1. **Detention:** Does the AI system influence decisions regarding detention conditions or treatment that could lead to inhuman or degrading treatment?
2. **External actions:** Does the AI system influence decisions regarding the requests of asylum seekers? Does the AI system apply to situations that reside in the EU's foreign policy and external actions?

v. **Art. 5 CFREU – Prohibition of slavery and forced labour (excluded from the horizontal application)**

- **Short explanation:**

Article 5 of the EU Charter of Fundamental Rights specifically addresses the prohibition of slavery, servitude, forced labour, and human trafficking, highlighting the EU and MS's commitment to combatting these severe human rights violations.

- **Guiding question:**

1. **Victim Identification:** Does the AI system have the capabilities to help identify victims of trafficking and forced labour? Does this status influence in any aspect the AI system?

vi. **Art. 6 CFREU – Right to liberty and security (excluded from the horizontal application)**

- **Short explanation:**

Article 6 of the EU Charter of Fundamental Rights ensures the right to liberty and security, which becomes particularly relevant in the domains of immigration, asylum, and criminal justice under the EU's Area of Freedom, Security and Justice. This article shapes the standards and practices concerning the detention and treatment of individuals within these

areas, ensuring that actions like arrests, detentions, or deportations comply with EU laws designed to protect personal freedom and security.

- **Guiding questions:**

1. **Impact on financial services:** Does the AI system use information about an individual's legal history to make decisions about creditworthiness or access to banking services? If so, how is this information weighed against other factors?
2. **Preventative Measures:** What safeguards and compliance measures are necessary to prevent the misuse of personal data that could lead to discrimination or unjust financial exclusion?

vii. Art. 7 CFREU – Right to privacy

- **Short explanation:**

Article 7 of the EU Charter of Fundamental Rights emphasizes the right to respect for private and family life. It acknowledges the broad and dynamic nature of family life within the EU, influenced by various factors including legislation on free movement, social security, and gender equality. The right to privacy under Article 7 ensures the protection of personal and family records against unauthorized or unnecessary collection, use, or disclosure of such information. This is particularly critical in contexts affecting children, gender equality, and migrants, where privacy and data protection are paramount.

- **Guiding questions:**

1. **Data Collection and Privacy:** Does the AI system collect personal data from family contexts? How does it ensure that such data collection respects the privacy and autonomy of family members?
2. **Impact on Children and Families:** How does the AI system address or potentially infringe upon the rights and protections specifically afforded to children under EU law?
3. **Influence on Gender Equality and Family:** Could the AI system inadvertently perpetuate stereotypes or biases that affect gender equality within family settings?

viii. Art. 8 CFREU – Right to data protection

- **Short explanation:**

Article 8 of the EU Charter of Fundamental Rights enshrines the right to the protection of personal data. It ensures that personal data must be processed fairly, for specified purposes, and on the basis of the consent of the person concerned or another legitimate basis laid down by law. Moreover, everyone has the right to access and rectify data collected about them. This right is particularly significant in the context of AI, where vast amounts of personal data can be processed automatically.

- **Guiding questions:**

1. **Data Accuracy and Integrity:** How does the AI system ensure the accuracy and integrity of personal data it processes?
2. **Limitation and Minimization:** Does the AI system restrict processing of personal data to the explicit purposes for which it was collected, and is the data collected minimally sufficient?
3. **Transparency and User Control:** How transparent is the AI system about data processing practices, and how does it empower users to control their personal data?

ix. **Art. 9 CFREU – Right to marry and found a family**

- **Short explanation:**

Article 9 of the EU Charter of Fundamental Rights addresses the right to marry and found a family. This right involves areas related to non-discrimination, cross-border recognition of family statuses, and public health services affecting family life.

- **Guiding questions:**

1. **Non-Discrimination:** How does the AI system ensure it does not discriminate based on sexual orientation, gender identity, or marital status in providing services related to family and marriage rights?
2. **Cross-Border Recognition:** Does the AI system have mechanisms to recognize and respect marriage and family statuses across EU member states, especially in managing data that crosses borders?
3. **Health Data Handling:** How does the AI system manage health data used (e.g.) in reproductive and family planning services to ensure compliance with EU health regulations and respect for privacy and consent?

x. **Art. 10 CFREU – Right to freedom of thought, conscience and religion**

- **Short explanation:**

Article 10 of the Charter of Fundamental Rights of the European Union guarantees the right to freedom of thought, conscience, and religion. This includes the right to hold beliefs, change religion or belief, and manifest one's religion or beliefs in worship, teaching, practice, and observance, either alone or in community with others. Areas of application include employment, public expression, education, healthcare.

- **Guiding questions:**

1. **Influence on Decision-Making:** Does the AI system potentially shape or influence individual decision-making in a way that limits the autonomy to make free choices, especially related to religious or philosophical beliefs? For example, do the system's recommendations or predictions steer someone away from their beliefs due to bias in data?

2. **Manifestation of Religious Beliefs:** Does the AI system accommodate or limit the expression of religious symbols or practices, either explicitly or through indirect bias? How does it handle decisions in sensitive areas, like healthcare or employment, where religious beliefs may affect participation?
3. **Religious Neutrality in Decision-Making:** Does the AI system remain neutral in scenarios where users' religious or non-religious beliefs come into play? Are there mechanisms to ensure that religious freedom is respected, such as ensuring that decisions regarding public-facing roles or policies on attire (e.g., religious headwear) are not biased?
4. **Conscientious Objection:** Does the AI system provide accommodations for individuals who might refuse to follow certain decisions or recommendations due to conscientious objections based on religious or ethical grounds? Can users override AI decisions in favour of their beliefs?

xi. Art. 11 CFREU – Right to freedom of expression

- **Short explanation:**

Freedom of expression constitutes a cornerstone of democratic societies. It is at the centre of individual autonomy, public discourse, and the functioning of the media. This right encompasses the freedom to hold opinions and to receive and impart information and ideas without undue interference. It also ensures the freedom and pluralism of the media which is at the core of a diverse and informed public sphere. This right impact areas such as: the right to hold opinions; the right to impart information and ideas; and the right to receive information and ideas.

- **Guiding questions:**

1. **Autonomy:** Does the AI system enhance or restrict an individual's ability to access, share, or express information and opinions? How does it impact the diversity of perspectives available to users?
2. **Limitation and Bias:** Does the AI system exhibit any biases that could lead to the suppression of certain viewpoints or the unfair promotion of others? Is there a risk of the system being used to censor or unfairly filter content?
3. **Media Pluralism:** How does the AI system affect the plurality of media sources and the diversity of content? Does it contribute to the concentration of media ownership or influence?

xii. Art. 12 CFREU – Right to assembly and association

- **Short explanation:**

The right to assembly and association guarantees individuals the freedom to gather peacefully and to form or join groups, including political parties, trade unions, and civic organisations. This right is essential for political participation, social solidarity, and the expression of collective interests. In this case, AI systems could limit the possibility of individuals assembling, associating, and participating in political and social life, for instance, by profiling individuals.

- **Guiding questions:**

1. **Assembly:** Does the AI system enhance or restrict the ability of individuals to organise and participate in assemblies or associations? How does it impact the ease with which people can connect with others with similar political, social, or economic interests?
2. **Surveillance:** Can the AI system potentially be used to monitor individuals' participation in assemblies or associations? Could it be misused to discourage or penalise involvement in such groups?
3. **Participation:** How does the AI system influence the functioning of associations? Does it support their role in expressing the political will of citizens?
4. **Bias in Association:** Does the AI system exhibit bias when considering individuals as part of an association, potentially affecting decisions like access to services or employment based on their group affiliations? How does the system mitigate risks of unfair treatment based on such biases?

xiii. Article 13 CFREU – Freedom of the arts and sciences

- **Short explanation:**

Article 13 of the CFREU deals with the freedom of the arts and sciences and aims to foster knowledge, culture, and innovation. It ensures that artistic expression and scientific research are free from undue restrictions. Academic freedom is also a key component, safeguarding the independence of educational institutions and the development of knowledge. AI systems raise primary questions about the creation of knowledge and the development of innovation.

- **Guiding questions:**

1. **Artistic and Scientific Freedom:** Does the AI system impact the ability of artists and researchers to explore new ideas and push boundaries without fear of limitation?
2. **Bias in knowledge creation:** Does the AI system exhibit biases that could lead to manipulating artistic or scientific content? Is there a risk of the system being used to foster or hide controversial or unconventional works?
3. **Impact on Academic Freedom:** How does the AI system influence the independence of academic institutions and researchers? Does it support the free exchange of ideas or introduce constraints that might hinder academic debate?

xiv. Article 14 CFREU – Right to education

- **Short explanation:**

Article 14 of the CFREU is about the protection of the right to education. It ensures access to quality education for everyone, regardless of their nationality or status. It includes the right to vocational training, non-discrimination in educational opportunities, and access to education for EU nationals and third-country nationals under certain conditions. The right to education also intersects with freedom of movement within the EU and the ability to pursue studies across member states.

- **Guiding questions:**

1. **Access to Education-Related Data:** Does the AI system take into account educational background when assessing creditworthiness, and how does it ensure that it does not disadvantage applicants from different educational systems or levels?
2. **Bias in Education-Based Credit Decisions:** How does the AI system avoid biases related to the type or level of education (e.g., vocational training vs. university degrees) when determining credit scores?
3. **Non-Discrimination in Educational Attainment:** Does the AI system account for educational attainment without introducing discrimination based on the location or quality of the institution attended?
4. **Support for Vocational and Lifelong Learning:** How does the AI system accommodate applicants who are pursuing further education or vocational training, particularly in terms of evaluating future creditworthiness?
5. **Educational Pathways for Migrants and Third-Country Nationals:** Does the system treat applicants fairly, regardless of whether their education was completed abroad or in a non-EU member state?

xv. **Art. 15 CFREU – Freedom to Choose an Occupation and Right to Engage in Work**

- **Short explanation:**

Article 15 of the EU Charter protects the right to engage in work and pursue a freely chosen occupation. This right encompasses equal treatment for EU citizens across Member States regarding employment opportunities, social, and tax advantages. For third-country nationals, once authorized to work, they are entitled to equal treatment in working conditions.

- **Guiding questions:**

1. **Employment Bias:** Does the AI system introduce biases related to the applicant's occupation or employment status? Does it favour specific job sectors or professions when determining creditworthiness?
2. **Job Mobility:** How does the system treat applicants who have worked across borders or in different countries within the EU? Is cross-border employment properly recognized without introducing penalties?
3. **Third-Country Nationals:** Are third-country nationals treated equitably compared to EU citizens when it comes to assessing creditworthiness based on employment data?
4. **Impact of Non-Traditional Work:** How does the AI system evaluate non-traditional work arrangements (e.g., freelance, gig economy) compared to standard full-time employment? Does it risk disadvantaging these applicants?
5. **Access to Employment-Based Credit Data:** How does the AI system ensure that individuals in lower-paying jobs or those with temporary contracts are not disproportionately penalized in credit evaluations?

xvi. **Art. 16 CFREU – Freedom to conduct a business**

- **Short explanation:**

Article 16 of the EU Charter protects the freedom to conduct business, including the right to start and manage economic activity. It is closely linked to the rights to property (Article 17) and work (Article 15).

- **Guiding questions:**

1. **Business Profiling:** Does the AI system profile individuals based on their business activities? Does it introduce biases against entrepreneurs in certain industries or sizes of businesses?
2. **Impact on Entrepreneurship:** How does the AI system treat self-employed or small business owners compared to individuals employed by larger organizations? Are there any inherent biases in assessing the creditworthiness of entrepreneurs?
3. **Business Continuity:** Could the AI system disproportionately affect the ability of business owners to access credit, thereby potentially impacting their ability to maintain or grow their business?
4. **Fairness:** Are there measures in place to ensure fairness in assessing applications from those running innovative or non-traditional businesses?
5. **Autonomy in Business Conduct:** Could decisions made by the AI system restrict an individual's ability to engage in economic activity or pursue entrepreneurial ventures? Does it take into account dynamic and evolving business environments?

xvii. **Art. 17 CFREU – Right to property**

- **Short explanation:**

Article 17(1) of the Charter protects the right to property, ensuring that individuals, companies, and entities are entitled to own, use, and dispose of their possessions. The right extends to both tangible and intangible assets and guarantees that property cannot be taken away except in cases of public interest and under fair compensation, which must be established by law. Article 17(2) specifically highlights the growing importance of intellectual property (IP) rights, including copyrights, patents, and trademarks. Intellectual property extends the scope of the right to property by ensuring that the creations of the mind, such as inventions, literary and artistic works, and symbols, are recognized and protected. Given the increasing reliance on knowledge-based economies, protecting intellectual property rights is crucial to fostering innovation, creativity, and economic competitiveness in the EU. In this context, EU law seeks to strike a balance between protecting IP rights and promoting the public interest, ensuring that the use of intellectual property does not unduly restrict innovation or fair competition.

- **Guiding questions:**

1. **Property Evaluation:** Does the AI system accurately assess an individual's or business's assets? Does it factor in intellectual property when determining creditworthiness?

2. **Bias in Property Assessment:** Are there biases in the AI system that could misvalue certain types of property or assets (e.g., small business assets vs. large corporation assets)?
3. **Impact on Asset Use:** Could decisions made by the AI system limit an individual's ability to leverage their property for loans or credit? Does it fairly evaluate properties and intellectual assets?
4. **Property Rights Protection:** Does the AI system safeguard property rights in its evaluations, ensuring that individuals are not unfairly penalized due to system errors or biases?

xviii. Art. 18 CFREU – Right to asylum [state based provision]

- **Short explanation:**

Article 18 of the Charter of Fundamental Rights of the European Union guarantees the right to asylum, framing it within the existing EU asylum system and protocols. This right ensures that individuals seeking international protection within the EU have access to a fair and functioning asylum process. The field of application of Article 18 is primarily centered on third-country nationals seeking protection within the EU.

- **Guiding questions:**

1. **Access to Credit:** How does the AI system assess third-country nationals seeking asylum in the EU when determining creditworthiness? Could their uncertain legal status impact access to financial services?

xix. Art. 19 CFREU – protection in the event of removal, expulsion or extradition [state based provision]

- **Short explanation:**

Article 19 of the EU Charter prohibits collective expulsion and protects individuals from being sent to a country where they may face torture, inhuman, or degrading treatment. It has two main components: Article 19(1) prevents the collective expulsion of EU and third-country nationals, protecting the right to individual consideration of cases; Article 19(2) prohibits extradition or removal of individuals to countries where they may face human rights abuses.

- **Guiding questions:**

1. **Exclusionary Practices:** Does the AI system used in credit decisions disproportionately exclude certain groups (e.g., third-country nationals)? Could this potentially constitute a form of indirect discrimination or collective expulsion from financial services?
2. **Risk of Bias:** Does the AI system consider the risk of individuals from certain regions being excluded from access to banking services, possibly reflecting biases based on nationality or asylum-seeking status?

xx. Art. 20 CFREU – Equality before the law

- **Short explanation:**

Article 20 of the EU Charter guarantees that “everyone is equal before the law.” This fundamental right is central to EU law, ensuring that all individuals receive fair and equal treatment in all situations where EU law applies. It plays a broad role in numerous areas, including employment, social welfare, taxation, and public services, where differences in treatment can arise.

- **Guiding questions:**

1. **Fairness in Credit Decisions:** Does the AI system treat all individuals equally, regardless of nationality, gender, age, or any other characteristic? Does it ensure that no discriminatory patterns are emerging in credit scoring?
2. **Data Inputs and Bias:** Are the data inputs or algorithms being used by the AI system free from any inherent biases that might lead to unequal treatment of certain groups or individuals?
3. **Transparency in Decision-Making:** Is the AI system transparent in how it determines creditworthiness? Can affected individuals access clear reasoning or justifications for decisions, ensuring they understand how their application is evaluated?

xxi. Art. 21 CFREU – Non discrimination:

- **Short explanation:**

- Article 21 of the EU Charter prohibits any discrimination based on nationality or other personal characteristics such as race, gender, religion, sexual orientation, and more. It is divided into two parts: one focusing on nationality discrimination (Art. 21(2)) and the other addressing broader status-based discrimination (Art. 21(1)). While Article 21(1) aligns with EU anti-discrimination laws, Article 21(2) specifically targets nationality discrimination within the scope of the Treaties.

- **Guiding questions:**

1. **Non-Discrimination in Credit Decisions:** Does the AI system treat applicants equally regardless of characteristics like nationality, ethnicity, gender, religion, or age? Are there safeguards against discriminatory patterns?
2. **Bias Mitigation:** Are there checks in place to ensure that the AI system does not unintentionally introduce bias based on personal characteristics like race or religion?
3. **Equal Access:** Does the AI system ensure that all individuals, regardless of nationality, have equal access to financial services and credit?
4. **Transparency in Data Usage:** Is there clarity on how personal data, especially sensitive attributes like nationality or genetic information, is used in making decisions about creditworthiness?

5. **Adherence to Anti-Discrimination Laws:** Does the system comply with both national and EU anti-discrimination laws when assessing credit applications?

xxii. Article 22 CFREU – Cultural, religious and linguistic diversity

- **Short explanation:**

Article 22 of the EU Charter emphasises that the European Union must respect and promote the diversity of cultures, religions, and languages across its member states. The recognition of this right underlines the commitment to the coexistence of various cultural identities based on the value of diversity, and ensuring that cultural, religious, or linguistic differences are not a basis for discrimination or exclusion. AI could fail to represent diversity and different cultural nuances.

- **Guiding questions:**

1. **Recognition of Diversity:** Does the AI system recognise diverse cultural, religious, and linguistic backgrounds when making decisions? Does it avoid imposing uniform criteria that may marginalise certain groups?
2. **Cultural Sensitivity in Decision-Making:** Are the algorithms sensitive to different cultural and religious practices or holidays, ensuring decisions are not biased against individuals due to cultural or religious observances (e.g., time off work for religious reasons)?
3. **Fair Representation of Diversity in Data:** Does the AI system's data set represent the cultural, religious, and linguistic diversity of the EU population, ensuring no group is underrepresented or misrepresented? How does it prevent cultural biases in data processing?

xxiii. Article 23 CFREU – Equality between women and men

- **Short explanation:**

Article 23 of the EU Charter mandates the equality between women and men. This right extends to different areas of life, including employment, work, and pay. It does not only promote gender equality but also allows for the implementation of positive actions (e.g., specific policies or measures) to benefit the under-represented sex, often women, in situations where historical or structural inequalities exist. AI decision-making could lead to gender-based discrimination.

- **Guiding questions:**

1. **Employment and Pay:** Does the AI system ensure that men and women are treated equally in employment decisions, promotions, and compensation? Are there mechanisms to prevent gender bias in job offers, performance reviews, or salary determination?
2. **Gender Bias and Stereotyping:** Are there safeguards against gender bias in the data used by the AI system? Does the system analyse patterns to detect and correct any discriminatory practices based on gender?

3. **Gender-specific issues:** Does the AI system account for gender-specific factors that might influence employment decisions, such as parental leave, caregiving responsibilities, or flexible working arrangements, ensuring that these do not discriminate against women?

xxiv. Article 24 CFREU – The Rights of the child

- **Short explanation:**

Article 24 of the EU Charter outlines children's rights to protection, care, and participation in decisions affecting them. It consists of three key principles. First, Children have the right to necessary protection and care for their well-being and freely express their views, which must be considered by age and maturity. Second, in all matters involving children, their best interests must be a primary consideration for both public authorities and private institutions. Third, every child has the right to maintain regular, personal relationships and contact with both parents, unless this would be contrary to their best interests.

- **Guiding questions:**

1. **Child Protection:** Does the AI system ensure that children's well-being is safeguarded in its services, such as online platforms or educational tools? Are there mechanisms to prevent harm to children, including data protection, safety from inappropriate content, and limits on access to age-restricted services?
2. **Children's Maturity:** Does the system allow children to interact with children and include their views, ensuring these are taken into account in decisions or services that concern them? Are the child's age and maturity level considered when gathering or responding to feedback from children?
3. **Best Interests:** Are the child's best interests a primary focus in any AI-driven decision-making processes, such as healthcare, education, or child protection services? Are there specific safeguards to prevent the system from prioritising other factors, such as financial gain, over the well-being of children?
4. **Harmful Content:** Does the system prevent access to content or interactions that could harm a child's development or well-being, such as violence, hate speech, or exploitation? How does it monitor and filter age-inappropriate content to protect young users?

xxv. Article 25 CFREU – The rights of the elderly

- **Short explanation:**

Article 25 of the EU Charter recognises the rights of elderly individuals to live with dignity, maintain their independence, and actively participate in social and cultural life. This article ensures that older people are treated with respect and provided with opportunities to engage fully in society, avoiding exclusion due to age.

- **Guiding questions:**

1. **Dignity of the Elderly:** Does the AI system ensure that the dignity of older adults is preserved when they engage with AI systems? Are there mechanisms to avoid ageist stereotypes in decision-making processes?
2. **Independence:** Does the AI system provide tools or features that promote the independence of the elderly, such as assistive technologies, mobility aids, or smart home devices? How does it ensure that elderly individuals can use the system effectively, considering potential barriers like limited tech skills?
3. **Participation:** Does the AI system create or support opportunities for older people to engage in social and cultural activities, either through online platforms, community-building tools, or access to cultural content? How does it encourage their active participation in society?
4. **Accessibility:** Is the system designed to be user-friendly for elderly users, with features like simplified interfaces, larger text, or voice-controlled commands? Are there efforts to reduce digital exclusion by making the technology accessible for people with reduced mobility or sensory impairments?

xxvi. Article 26 CFREU – Integration of Persons with Disabilities

- **Short explanation:**

Article 26 of the EU Charter recognises the right of persons with disabilities to benefit from measures that ensure their independence, social and occupational integration, and active participation in community life. This article underscores the importance of creating an inclusive environment where individuals with disabilities have equal opportunities to engage in all aspects of society.

- **Guiding questions:**

1. **Independence:** Does the AI system include features or tools that enhance the independence of persons with disabilities, such as assistive technologies, adaptive interfaces, or personalised support? How does it ensure these tools are effective and accessible?
2. **Integration:** Does the AI system support the social and occupational integration of individuals with disabilities, for example, by facilitating access to employment opportunities, social activities, or educational resources? Are there features that help bridge gaps and foster inclusion in various settings?
3. **Avoidance of Discrimination:** Does the AI system ensure that persons with disabilities are not subjected to discrimination or bias in its processes? Are there safeguards in place to detect and address potential biases?

xxvii. Art. 27 CFREU – Workers' rights:

- **Short explanation:**

Article 27 of the EU Charter establishes the right of workers to information and consultation within their organizations. This right is critical in EU labor law, ensuring that employees or their representatives are informed and consulted on matters affecting the workplace. The scope of this right is dependent on national laws and EU Directives, and it typically applies

at various levels such as establishments, undertakings, and company groups. However, Article 27 lacks direct effect unless implemented through specific national or EU legislation.

- **Guiding questions:**

1. **Worker Consultation:** Does the AI system inform and consult workers (or their representatives) regarding decisions impacting their employment or the workplace environment?

xxviii. Art. 28 CFREU – Right to collective actions:

- **Short explanation:**

Article 28 of the EU Charter guarantees the right of workers and their representatives to engage in collective bargaining and take collective action, including strike action, in defense of their interests. This right applies to the institutions, bodies, offices, and agencies of the EU, as well as to Member States when implementing EU law. However, this right must be exercised in accordance with both Union and national laws, making it subject to limitations based on public interest or economic rights.

- **Guiding questions:**

1. **Collective Impact:** Does the AI system influence workers' rights to participate in collective bargaining or collective actions like strikes?

xxix. Art. 29 CFREU – Right of access to placement services:

- **Short explanation:**

Article 29 of the Charter guarantees the right to access free placement services, facilitating labor mobility and helping individuals find employment within the EU. This right is closely tied to the freedom of movement for workers, ensuring that both EU nationals and workers from member states can benefit from equal access to job opportunities. The role of national and European employment services, including platforms like EURES, is essential in coordinating job vacancies, applications, and labor market data, contributing to a balanced, integrated labor market.

- **Guiding questions:**

1. **Job Placement:** How does the AI system consider access to free placement services when evaluating the creditworthiness of individuals, especially those transitioning between jobs or relocating for employment?
2. **Worker Mobility:** Does the system account for potential biases or barriers related to worker mobility, such as nationality, in determining creditworthiness?

3. **Transparency:** Does the AI system transparently utilize job-related data (such as employment status or job search activities) fairly to assess an individual's financial stability and creditworthiness?
4. **Transitions:** How does the system handle individuals undergoing job transitions or relying on employment services? Does it offer flexibility to accommodate temporary unemployment or job-seeking periods?

xxx. Art. 30 CFREU – Protection in the event of unjustified dismissal:

- **Short explanation:**
Article 30 of the Charter protects workers from unjustified dismissal. This right is meant to prevent arbitrary terminations, ensuring that any dismissal follows a lawful procedure, includes appropriate reasons, and provides redress, such as compensation or dispute resolution mechanisms.
- **Guiding questions:**
 1. **Employment Stability:** How does the AI system account for the risk of unjust dismissal when assessing an individual's creditworthiness, especially in cases of sudden employment termination?

xxxi. Art. 31 CFREU – Right to fair and just work conditions

- **Short explanation:**
Article 31 of the Charter of Fundamental Rights of the EU ensures the right to fair and just working conditions. This includes protecting workers' safety, health, and dignity in their working environment. It also guarantees that workers receive adequate rest and paid annual leave.
- **Guiding questions:**
 1. **Impact of Employment Status:** How does the AI system account for temporary or precarious employment and its potential effect on a worker's health and well-being when assessing financial risk?

xxxii. Art. 32 CFREU – Prohibition of child labour

- **Short explanation:**
Article 32 of the Charter of Fundamental Rights of the EU ensures the protection of children from economic exploitation and harmful labor practices. It restricts child labor and ensures that any work undertaken by young people is done under safe conditions. It aims to balance the need for vocational opportunities with educational development and safeguarding the health, safety, and well-being of minors.

- **Guiding questions:**

1. **Protection of Young Workers:** Does the AI system flag applicants under the minimum working age or those with limited work history due to youth protections?

xxxiii. Art. 33 CFREU – Family and professional life:

- **Short explanation:**

Article 33 ensures protection for family life and work-life balance, emphasizing social and economic protection for families.

- **Guiding questions:**

1. **Maternity/Parental Leave:** How does the AI consider applicants who may be on maternity or parental leave? Does it flag these circumstances as high-risk unjustly, or does it account for the temporary nature of such absences?

xxxiv. Art. 34 CFREU – Social security and social assistance:

- **Short explanation:**

Article 34 of the Charter ensures access to social security, housing, and social assistance. It covers citizens and residents within the EU and supports their right to receive appropriate support, especially during vulnerable times, such as unemployment or illness.

- **Guiding questions:**

1. **Social Security Impact:** Does the AI system consider applicants' access to social security benefits, and how might a lack of access influence credit decisions?
2. **Vulnerability Assessment:** How does the AI handle applicants experiencing economic vulnerability due to illness, unemployment, or similar situations, and does it account for assistance or benefits they receive?
3. **Non-Discrimination:** Does the AI ensure no indirect discrimination against applicants who rely on social or housing assistance when evaluating their creditworthiness?

xxxv. Art. 35 CFREU – The right to healthcare

- **Short explanation:**

Article 35 of the Charter ensures access to preventive health care and medical treatment under conditions established by national laws and practices. It mandates that a high level of human health protection be integrated into all Union policies and activities.

- **Guiding questions:**

1. **Health-Related Expenditures:** Does the AI system take into account the applicant's medical costs or health-related financial burdens when determining creditworthiness?
2. **Discrimination:** Could the AI system disadvantage applicants with health conditions, leading to potential indirect discrimination?
3. **Access to Care:** Does the AI ensure that applicants, especially those in economic distress due to medical conditions, are not unfairly assessed or denied access to credit, which could hinder their ability to access necessary medical care?
4. **Vulnerability Consideration:** How does the AI evaluate applicants who are in economically vulnerable positions due to health issues, and does it consider available health care support systems in its assessment?

xxxvi. Art. 36 CFREU – Access to service of general economic interest [state based provision]

- **Short explanation:**

Article 36 of the Charter concerns access to services of general economic interest (SGEI), ensuring that everyone has the right to access these services, especially those essential for maintaining dignity and ensuring welfare (e.g., electricity, water, healthcare).

- **Guiding questions:**

1. **Impact on Vulnerable Groups:** Could the AI system indirectly disadvantage vulnerable populations (e.g., Roma, economically marginalized) in accessing essential services through its creditworthiness assessment?

xxxvii. Art. 37 CFREU – Environmental protection [state based provision]

- **Short explanation:**

Article 37 of the Charter focuses on integrating environmental protection into all Union policies, aiming for a high level of protection and improvement in environmental quality. However, it does not establish any specific individual right to environmental protection. Instead, it serves as a guiding principle for EU institutions and Member States when implementing Union law, especially concerning policies on climate, biodiversity, pollution, and waste management.

- **Guiding questions:**

1. **Environmental Impact:** Does the AI system consider environmental factors in its assessment, such as the impact of climate risks on creditworthiness?
2. **Sustainability Incentives:** Does the AI system factor in potential benefits or risks associated with green initiatives or compliance with environmental laws when evaluating clients?

xxxviii. Art. 38 CFREU – Consumer protection

- **Short explanation:**

Article 38 of the Charter of Fundamental Rights of the European Union ensures that all Union policies and activities provide a high level of consumer protection. It aligns closely with Article 169 TFEU, which promotes consumer health, safety, and economic interests. The article has broad relevance, influencing not only consumer protection but also harmonizing laws affecting trade, free movement, and public policies across the Union.

- **Guiding questions:**

1. **Consumer Transparency:** Does the AI system ensure transparency in its assessments, so consumers understand why credit decisions are made?
2. **Consumer Protection:** How does the system ensure fairness in evaluating consumer creditworthiness, particularly for vulnerable consumers?
3. **Data Usage:** How does the system ensure that consumer data is protected, and does it comply with consumer protection laws and standards?
4. **Dispute Resolution:** What mechanisms are in place for consumers to challenge decisions made by the AI, ensuring their right to a remedy?
5. **Bias Mitigation:** Does the AI consider and mitigate biases that could harm certain consumer groups, ensuring nondiscriminatory practices in evaluating creditworthiness?

xxxix. Art. 39 CFREU – The right to vote [state based provision]

- **Short explanation:**

Article 39 of the Charter of Fundamental Rights of the EU secures the right for citizens of the Union to vote and stand as candidates in elections to the European Parliament. This right applies to EU institutions and Member States implementing Union law, particularly under the Act on Direct Elections and Council Directive 93/109/EC. It relates closely to Article 40 (municipal election rights) and promotes democratic participation within the EU's political processes.

- **Guiding questions:**

1. **Public Profile Impact:** How does the public profile and political activities of the candidate (e.g., income from political roles, endorsements, or campaign contributions) affect their financial stability and creditworthiness?
2. **Cross-Border Considerations:** If the candidate has political ties or activities across multiple EU Member States, how is their financial data collected, and does the system account for variations in credit assessment standards across borders?

xl. Art. 40 CFREU – The right to vote and to stand as a candidate

- **Short explanation:**

Article 40 guarantees EU citizens the right to vote and stand as a candidate in municipal elections, regardless of their nationality within the Union. This right reflects the EU's commitment to democratic participation and equal treatment, aligning with broader Union

values. However, it applies only to local elections, and does not extend to national elections. It ensures non-discriminatory access to political participation at the municipal level for EU citizens residing in other Member States.

- **Guiding questions:**

1. **Political Contributions:** Has the candidate declared all political contributions and fundraising activities, and are there any irregularities?
2. **Public Records:** Do public records reveal any legal or financial actions against the candidate that might indicate financial irresponsibility?
3. **Conflict of Interest:** Are there any potential conflicts of interest between the candidate's financial dealings and their political role?
4. **Income Stability:** Is the candidate's income primarily tied to their political career, and if so, how stable is that source of income considering electoral outcomes and term limits?
5. **Financial Transparency:** Has the candidate demonstrated financial transparency in line with public office expectations, including disclosure of assets, debts, and financial interests?
6. **Potential Conflicts of Interest:** Are there any potential conflicts of interest between the candidate's financial obligations and their public role that may affect their credit risk profile?
7. **Reputation Risk:** How does the candidate's public reputation, including any legal or ethical concerns, affect their perceived reliability as a borrower or credit applicant?
8. **Electoral Campaign Finances:** How do campaign financing laws (e.g., donations and spending limits) impact the candidate's overall financial profile and ability to service personal debt?

xli. Art. 41 CFREU – The right to good administration [state based provision]

- **Short explanation:**

Article 41 of the EU Charter guarantees the right to good administration, ensuring that individuals have the right to fair treatment in dealings with the EU institutions. This includes the right to be heard, access to one's file, and receiving reasons for decisions. While the Article specifically addresses EU institutions, bodies, offices, and agencies, the general principle of good administration also applies to Member States when acting within the scope of EU law, as clarified by case law.

- **Guiding questions:**

1. **Transparency in Decision-Making:** How does the AI system ensure transparency in the credit evaluation process?
2. **Right to Be Heard:** Does the system allow individuals to provide input or clarify their financial data before decisions are made?

xlii. Art. 42 CFREU – The right to access documents

- **Short explanation:**

Article 42 of the Charter grants every citizen of the Union the right to access documents of EU institutions, bodies, offices, and agencies. This right is closely linked to Article 15(3) TFEU and is governed by secondary legislation, most notably Regulation 1049/2001.

- **Guiding questions:**

1. **Access to Personal Data:** Can individuals access all the financial documents and data the AI used to assess their creditworthiness?
2. **Transparency of Evaluation Criteria:** How transparent is the bank about the criteria used by the AI to make credit decisions?
3. **Request for Information:** Is there a system in place for individuals to request additional information or clarification on how their credit score was determined?

xl.iii. Art. 43 CFREU – European Ombudsman [state based provision]

- **Short explanation:**

Article 43 of the Charter guarantees EU citizens and residents the right to complain to the European Ombudsman. This right applies to situations where EU institutions, bodies, offices, or agencies have committed maladministration. If a bank's AI system evaluating creditworthiness is part of an EU-regulated institution, individuals could potentially file a complaint with the Ombudsman if the evaluation process appears unfair or lacks transparency. The Ombudsman ensures that EU bodies adhere to principles of good administration.

- **Guiding questions:**

1. **Filing a Complaint:** Can individuals who feel unfairly assessed by the bank's AI system in terms of creditworthiness file a complaint with the European Ombudsman?
2. **Ombudsman's Role:** How does the European Ombudsman address complaints related to credit evaluation systems used by banks, particularly when AI is involved?
3. **Institutional Oversight:** What powers does the Ombudsman have to investigate potential maladministration in AI-driven credit assessments by financial institutions?
4. **Complaint Procedure:** What is the process for filing a complaint with the Ombudsman regarding the transparency and fairness of AI-based credit assessments?

xliv. Art. 44 CFREU – The right to petition [state based provision]

- **Short explanation:**

Article 44 of the Charter provides the right for every citizen or resident of the EU to petition the European Parliament on matters within the Union's scope of activity. If a bank uses AI to evaluate creditworthiness in a way that violates EU laws or principles, individuals can petition the European Parliament to review the matter. This right serves as an avenue for individuals to raise concerns about transparency or fairness in credit evaluation within the Union's legal framework.

- **Guiding questions:**

1. **Right to Be Heard:** How can the AI credit evaluation system be designed to address consumer petitions regarding inaccurate decisions? How could AI systems impact a person's ability to petition on financial issues related to credit access?

xliv. Art. 45 CFREU – Freedom of movement and residence

- **Short explanation:**

Article 45 of the Charter grants EU citizens the right to move and reside freely within the territory of the Member States. This right mirrors the provisions of Article 20(2)(a) TFEU and is fundamental to the European integration project. It primarily emphasizes that free movement rights, although central, must align with conditions defined in EU Treaties.

- **Guiding questions:**

1. **Non discrimination:** Does the AI system ensure non-discriminatory access to credit for EU citizens exercising their right to free movement?
2. **Mobile citizens:** How does the bank's credit evaluation process address the needs of mobile citizens, such as those frequently moving between EU countries?
3. **Data sources:** Are the data sources for AI credit evaluations standardized across different EU Member States, ensuring fairness for citizens residing in different regions?

xlvi. Art. 46 CFREU – Diplomatic and consular protection [state based provision]

- **Short explanation:**

Article 46 of the Charter grants EU citizens the right to diplomatic and consular protection from any Member State's embassy or consulate when outside the EU and their own country lacks representation. This provision supports the practical application of EU citizenship and fosters European solidarity. It is directly linked to Article 20 and Article 23 TFEU, which outline similar rights and confer powers for the necessary implementation of protection measures abroad.

- **Guiding questions:**

1. **Diplomatic or consular services:** Are diplomatic or consular services used as part of the verification process for citizens residing abroad?

xlvii. Art. 47 CFREU – Fair trial [state based provision]

- **Short explanation:**

Article 47 of the EU Charter of Fundamental Rights is centered on the right to an effective remedy and to a fair trial. It encompasses the principles of access to justice, ensuring that any person whose rights under EU law are violated has the right to a fair and just legal remedy.

This includes the necessity for transparency and clarity in any judicial or quasi-judicial process.

- **Guiding questions:**

1. **Transparency of Decision-Making:** Does the AI system adequately inform users that a decision is derived from an algorithm? What measures are in place to ensure users understand this?
2. **Reliability Communication:** Does the AI system provide information on the reliability of its outcomes, such as error rates or confidence intervals?
3. **Reversal of Burden of Proof:** Does the use of the AI system in any way imply a reversal of the burden of proof that could prejudice the defense's ability to contest the evidence?
4. **Explainability of Decisions:** Is there a clear explanation accessible to users about how and why decisions were made by the AI system?
5. **Influence on Decision-Making:** Is it clear to what extent the AI system's output influences overall decisions in a legal or quasi-legal context?
6. **Redress Mechanisms:** Are there adequate mechanisms in place for redress or appeal if the AI system causes harm or an adverse outcome?

xlvi. Art. 48 CFREU – Presumption of innocence and right to defense [state based provision]

- **Short explanation:**

Article 48 of the Charter guarantees the presumption of innocence until proven guilty in legal proceedings. It mirrors Article 6(2) of the European Convention on Human Rights (ECHR) and applies both in criminal law and in procedures where severe sanctions are imposed. This principle also applies to EU competition law and other legal proceedings involving sanctions. The right ensures that the burden of proof lies with the prosecuting authority and promotes fairness in judicial decisions.

- **Guiding questions:**

1. **Presumption of Innocence:** Does the AI system avoid prejudging an applicant's creditworthiness based on incomplete or past data?
2. **Transparency:** Can the applicant access the data used to assess their creditworthiness and challenge any inaccuracies?
3. **Right to Defense:** Is there an avenue for applicants to appeal or contest the AI's assessment of their creditworthiness?
4. **Fair Treatment:** How does the AI ensure the equal treatment of all applicants regardless of their previous credit history or background?

xli. Art. 49 CFREU – Principle of legality and proportionality of criminal offences and penalties [state based provision]

- **Short explanation:**

Article 49 of the Charter protects the principle that criminal offenses and penalties must be clearly defined by law (legality) and must not be excessive in relation to the offense (proportionality). These principles limit both national and EU powers to create or enforce criminal laws, ensuring fairness throughout the legislative, interpretative, and enforcement stages of EU and national law.

- **Guiding questions:**

1. **Legality:** Are the rules for assessing creditworthiness clear and accessible to applicants?
2. **Proportionality:** Is the assessment of an applicant's credit risk proportionate to the financial data used, avoiding overly punitive decisions?

1. **Art. 50 CFREU – *ne bis in idem* – the right to not be tried or punished twice [state based provision]**

- **Short explanation:**

Article 50 of the Charter enshrines the principle of *ne bis in idem*, preventing an individual from being tried or punished again for the same offence within the EU after a final verdict. This principle ensures fairness in criminal proceedings, safeguards finality of legal judgments, and applies broadly across the EU's criminal law framework and administrative sanctions resembling criminal law.

- **Guiding questions:**

1. **Consistency in Decision-Making:** Does the AI system ensure that applicants are not penalized multiple times for the same financial misstep?
2. **Transparency:** Are the criteria and decisions of the AI clearly explained to ensure there is no double evaluation or punishment?
3. **Review Mechanisms:** How does the system prevent repetitive evaluations of an applicant's creditworthiness for the same issue?