FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

ASSESSING HIGH-RISK ARTIFICIAL INTELLIGENCE

FUNDAMENTAL RIGHTS RISKS

REPORT

FRA

Assessing High-risk Artificial Intelligence: Fundamental Rights Risks

Vienna, 2025

Luxembourg: Publications Office of the European Union, 2025

PRINT	ISBN 978-92-9489-678-0	doi:10.2811/0062513	TK-01-25-028-EN-C
PDF	ISBN 978-92-9489-677-3	doi:10.2811/3952332	TK-01-25-028-EN-N

Photo credits:

Cover: © Putilov_denis / Adobe Stock
Page 11: © Nina Lawrenson / peopleimages.com / Adobe Stock
Page 15: © Gorodenkoff / Adobe Stock
Page 17: © Qunica / Adobe Stock
Page 19: © InfiniteFlow / Adobe Stock
Page 21: © Framestock / Adobe Stock
Page 26: © Kiattisak / Adobe Stock
Page 31: © Olivier Le Moal / Adobe Stock
Page 34: © Greentech / Adobe Stock
Page 40: © Prostock-studio / Adobe Stock
Page 45: © Cultura Creative / Adobe Stock
Page 53: © Nattawit / Adobe Stock
Page 57: © Wpw / Adobe Stock

Foreword

The AI Act – the world's first comprehensive AI legislation – demonstrates EU leadership in the rapidly evolving field of artificial intelligence. In 2025, the European Commission announced the world's largest public investment in AI to date, recognising the significant role it plays in our daily lives and economies.

But with AI comes risks and responsibilities. FRA's research shows that AI systems can profoundly impact fundamental rights – especially when used for decision-making in sensitive areas like who receives public benefits or who gets asylum. In these so-called 'high-risk' systems or areas, AI must be used with utmost caution and accountability.

Our report highlights that certain provisions in the AI Act may be interpreted in ways that limit the protection of fundamental rights. Without further guidance on the classification of 'high-risk AI-systems', the application of the AI Act could lead to loopholes in practice. Our findings are based on interviews with high-risk AI providers, deployers and experts in the field. Many of those developing, selling and using AI in high-risk areas do not know how to systematically assess or mitigate risks to fundamental rights. Awareness of the different rights affected remains low.

Assessing fundamental rights when using AI is not only good practice; it leads to better-performing technology. In 2025, we have heard many calls for simplification of legislative requirements to boost European innovation and competitiveness. Rights-compliant rules are essential to build public trust and support responsible uptake. If implemented as intended, the AI Act can foster innovation and provide legal certainty for businesses. It is therefore also key to continue to focus efforts on the timely implementation of the AI Act.

In November, the European Commission published a proposal for a Digital Omnibus on AI Regulation with amendments aiming to support competitiveness. Changes to the AI Act should not jeopardise fundamental rights protection. Regardless of potential delays, Member States, providers and deployers of AI should start preparing for the Act's high-risk requirements to enter into force, as the checks and balances included in the AI Act support the development of high-quality AI.

A fundamental rights approach to AI is not only feasible but directly supports the trust needed for innovation and competitiveness. It helps deliver on the EU's commitment to innovate without compromising its own values. As the world adapts to a new way of living with AI, there is an opportunity to show leadership and demonstrate that progress can go hand in hand with protecting fundamental rights.

Sirpa Rautio
Director

Contents

Key findings and FRA opinions

Artificial intelligence (AI) continues to be high on the policy and political agenda in 2025. The term captures a variety of technological developments, mainly based on the use of data to make predictions or create desired outputs, such as risk scores, images or text.

It is well known that the use of AI does not come without risks, especially to fundamental rights. The use of AI may reveal private information about people and can put vulnerable groups at further disadvantage. It may also be used without fully understanding its risks, which in turn presents challenges for remedying any adverse impacts. Most notably, such risks vary depending on the area and context of AI use.

To react to such risks and promote the development and use of trustworthy AI, the EU adopted the Artificial Intelligence Act (AI Act) in 2024. As an EU regulation, it is directly applicable in the EU Member States. One of the purposes of the AI Act is to ensure a high level of fundamental rights protection, which is enabled through several of its provisions. This report focuses on the key provisions of the AI Act and how it can be used for effective fundamental rights protection.

The AI Act

The AI Act is the first binding regional legal framework on AI. It entered into force on 1 August 2024. It has two interrelated goals: (1) to improve the functioning of the internal market, support innovation and promote the uptake of human-centric and trustworthy AI and (2) to ensure a high level of protection of health, safety and fundamental rights – including democracy, the rule of law and environmental protection – against the harmful effects of AI systems.

The AI Act follows a risk-based approach to regulating AI, distinguishing between (1) prohibited AI practices that pose unacceptable risks, (2) high-risk AI systems to which the majority of the AI Act's requirements pertain, (3) limited-risk AI systems with specific transparency requirements and (4) minimal-risk AI systems falling outside the scope of the AI Act. In addition, the act contains specific rules for the development of general-purpose AI models, which can be used for a wide range of distinct tasks.

Source: **Regulation (EU) 2024/1689 (AI Act)**.

The analysis in the report is based on interviews with those developing and selling AI ('providers') and those using AI for specific purposes ('deployers'). The report focuses on selected use cases in the areas that are considered high-risk areas under the AI Act, namely those of asylum, education, employment, law enforcement and public benefits (the final category is considered to broadly include essential public assistance benefits and services, such as social security benefits, social services and assistance, and housing). These are complemented by a small number of focus groups and interviews with rights holders about selected AI use cases in the areas of education, employment and law enforcement.

The interview period, between summer 2024 and early 2025, coincided with the early days after the adoption of the AI Act. This is reflected in interviewees' responses, as many were awaiting further guidance. Overall, while several interviewees identify a number of challenges with the implementation of the AI Act, they also highlight opportunities. These opportunities relate to what the AI Act can do for fundamental rights protection and responsible innovation in practice, including creating a level playing field between providers.

To ensure effective fundamental rights protection and increase certainty about the AI Act's application, the definition of an AI system should be interpreted broadly.

Ensure a broad interpretation of the definition of an artificial intelligence system

The AI Act applies to systems that meet the definition of an 'AI system' contained in Article 3(1) of the AI Act.

'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

Article 3(1) of the AI Act

FRA OPINION 1

With respect to the use of AI, to ensure the effective protection of fundamental rights (as enshrined in the Charter of Fundamental Rights of the European Union (the Charter)) and to increase certainty about the AI Act's application, the definition of an AI system in Article 3(1) of the AI Act should be interpreted broadly, covering systems with lower levels of complexity and automation. The Commission should encourage authorities in charge of the implementation of the AI Act (e.g. market surveillance authorities) to apply a broad interpretation of the definition when providing further guidance to providers and deployers of AI.

While this definition is phrased broadly, it also leaves room for interpretation, as highlighted by some of the interviewees during this research. This could lead to the exclusion of certain, less complex, systems from the scope of the AI Act – for example systems that use logistic regression. As demonstrated in the **European Union Agency for Fundamental Rights (FRA) 2022 report *Bias in Algorithms***, such systems can still contain biases and lead to discrimination.

Less complex systems could still have a profound impact on fundamental rights when used in high-risk areas. Some concerns have been raised over the possibility of developers circumventing the requirements of the AI Act by reclassifying AI systems as traditional software, as shown in the **responses** to the European Commission's public consultation on the definition of an AI system. These concerns are also reflected in some of the interviews done for this research.

In 2025, after the fieldwork for this project was concluded, the European Commission published guidelines on the definition of an AI system established by the AI Act (**C(2025) 924 final**), which provide some further clarity on the matter. The guidelines exclude from the scope of the definition established methods that have been used for many years to improve the efficiency of optimisation algorithms used in computational problems.

The description of AI, as set out in the Commission guidelines, remains open to interpretation and has been subject to criticism. The exclusion of 'simpler' systems fails to consider that such systems can be as performant as more complex systems. Ultimately, these 'simpler' systems can also have an adverse impact on fundamental rights if checks and balances are not in place. A wide and inclusive way of applying the law will not only increase the protection of fundamental rights but also enhance legal certainty and the quality of products and services.

The AI Act's filter for high-risk systems must be applied clearly and narrowly to prevent loopholes that could undermine fundamental rights protection across the EU.

Avoid 'filter' loopholes with respect to fundamental rights compliance

The AI Act uses a risk-based approach, meaning that its requirements apply only to the development and use of certain AI systems, most notably high-risk AI systems as defined in Article 6 of the AI Act. These are products (or safety components of products) covered by the EU product safety legislation listed in Annex I to the AI Act and requiring a third-party conformity assessment. This means that certain products need to be evaluated by organisations that are not the manufacturer or provider. In addition, the AI Act covers certain systems used in high-risk areas that can impact people's lives, such as education, employment or law enforcement, as listed in Annex III to the AI Act.

Some AI systems that generally fall under the high-risk AI use cases listed in Annex III to the AI Act may nevertheless be excluded from meeting the requirements for high-risk AI if they do not pose a significant risk of harm to people's health, safety or fundamental rights. This is stipulated in a derogation clause – also known as the 'filter' – in Article 6(3) of the AI Act. This covers cases in which the use of the system does not materially influence the outcome of decision-making. The filter applies if any of the four conditions listed in Article 6(3) of the AI Act are met, among which is an AI system being intended to perform 'a narrow procedural task' or 'a preparatory task'. Providers deciding to apply the filter have to register this decision in an EU database pursuant to Article 6(4) of the AI Act, meaning that there will be an overview of filter applications in practice. Certain exceptions limiting the transparency of these decisions apply, however, in the areas of law enforcement, migration, asylum and border control management (Article 49(4) of the AI Act).

Some experts are concerned about too broad an interpretation of the filter. Some respondents also highlight the inherent high-risk nature of systems used in certain areas, such as recruitment or public benefits. While the initial mapping conducted as part of this research suggested that the systems explored could be high risk, interviewees do not always agree with this assessment or are not sure about it. Some interviewees think that their system falls under the filter. In some cases, respondents referring to the same system have different views on its classification. This highlights the different interpretations of what constitutes a high-risk AI system, with possible reasons for this mismatch including interviewees' background knowledge about the system and about the AI Act. As the application of the filter will depend on providers' self-assessments, this could lead to misclassifications in practice.

FRA OPINION 2

The Commission and the AI Board should provide a clear and narrow understanding of the filter set out in Article 6(3) of the AI Act when developing the guidelines on the classification rules for high-risk AI systems. A narrow interpretation is needed, as the use of AI systems in areas listed in Annex III can have an impact on people's lives and seemingly supportive technologies can still significantly influence decision-making.

FRA OPINION 3

The Commission and the national competent authorities should proactively and carefully monitor the practical application of the filter. They should do so by collecting evidence from different sources, including through the EU and national databases for registering filter applications and through the AI Board. Because of their limited registration obligations, particular attention should be paid to systems used in the areas of law enforcement, migration, asylum and border control management. Should providers interpret the filter in a contradictory manner or in too broad a manner, thereby unduly limiting the protection of fundamental rights in practice, the Commission should consider amending Article 6(3) by deleting any of the conditions laid out therein, in line with its powers under Article 6(7) of the AI Act.

In terms of filter criteria, it is important to note that tasks that are considered preparatory can still, in some cases, influence decision-making, such as information selection or categorisation. In these cases, errors and biases in AI systems may create risks to fundamental rights. For example, a system used for language assessments to support country-of-origin determination in asylum procedures may be classified as performing a preparatory task. However, if this system's assessment is wrong (e.g. because of biases in the system) and the assessment is not corrected by those working with the system's outcomes, this could present a risk to the right to asylum.

There is the potential for the filter to substantially limit the number of AI systems held to the high-risk requirements in the AI Act, despite such systems being deployed in high-risk areas and posing potential risks to fundamental rights. As the application of the filter will be determined by providers' self-assessments, it is crucial that the rules on classification are interpreted in a uniform manner that ensures a high level of fundamental rights protection across the EU. There is a risk that providers could use the filter to circumvent the necessary safeguards.

The European Commission has been tasked with developing – after consulting the European AI Board, which comprises Member States' representatives – guidelines on the classification of AI systems as high risk (Article 6(5) of the AI Act). Article 6(6) and (7) of the AI Act sets out that the Commission can change or delete any of the filter conditions through delegated acts. According to Article 6(7) of the AI Act, filter conditions can be deleted 'where there is concrete and reliable evidence that this is necessary to maintain the level of protection of health, safety and fundamental rights'.

Clear and consistent guidance is needed to ensure that risk assessments for high-risk AI systems effectively protect all fundamental rights. Identifying and mitigating fundamental rights risks will promote responsible innovation and support fair competition by helping providers create better and more trustworthy AI.

Guide providers and deployers to use fundamental rights impact assessments

FRA OPINION 4

Guidance related to FRIAs under the standards under Article 9 and the FRIA template under Article 27 of the AI Act should cover the main cross-cutting fundamental rights concerns, namely privacy and data protection, equality and non-discrimination, and access to effective remedies. In addition, the impact on other rights should be considered and set out in accompanying guidance. The guidance should be tailored to specific high-risk areas and should provide examples of how high-risk AI systems might have an impact on certain rights in a practical and understandable manner.

FRA OPINION 5

The Commission and Member States are encouraged to provide guidance on how the AI Act can be best implemented in practice. Any simplification considerations need to be based on evidence and informed by the experiences of various stakeholders involved in the AI Act's implementation. They must not lower existing fundamental rights protection.

Under the AI Act, providers and certain deployers of high-risk AI systems will have to assess their systems' risks to fundamental rights. Providers will have to do so as part of their risk management system under Article 9 of the AI Act, the requirements for which will be developed further through standards. Certain deployers of high-risk AI systems listed in Annex III to the AI Act, except for those deploying systems in the area of critical infrastructure, will have to conduct fundamental rights impact assessments (FRIAs) under Article 27 of the AI Act. At the time of publishing this report, the European AI Office was in the process of developing the FRIA template, with FRA's assistance.

These assessments and the accompanying guidance were not yet applicable when conducting this research. This is reflected in the findings of this report, as several respondents across different high-risk areas express a certain level of uncertainty concerning the AI Act's interpretation and have concerns about how regulatory obligations could be met in practice. Nevertheless, there is also a definite undercurrent of optimism among many respondents about what the AI Act could mean for fundamental rights protection and responsible innovation in practice. Several respondents indicate that the regulation helps providers to create better AI by helping them to more effectively assess AI risks and by creating a level playing field between providers. This is especially important to note in the light of ongoing discussions on the simplification of existing laws, including in the digital area.

None of the providers or deployers of potential high-risk AI systems interviewed conducts risk assessments that take fundamental rights into account in a structured manner. Current risk assessment practices focus mostly on data protection, technical aspects or legal and/or business risks. Some providers and deployers also investigate bias in relation to non-discrimination.

This largely reflects findings from **FRA's 2020 report *Getting the Future Right – Artificial intelligence and fundamental rights***, which showed that there is a focus on technical and data protection (impact) assessments in practice. While the interviewees of that report showed awareness about the potential impact of AI systems on the rights to privacy and data protection (Articles 7 and 8 of the Charter), and to some extent non-discrimination (Article 21 of the Charter), their awareness about the impact on other rights was limited. Compared with the situation in 2020, there is a positive trend of a higher level of awareness about non-discrimination issues linked to AI among the providers and deployers interviewed for this report. However, this level of awareness does not necessarily translate into better assessments and mitigation measures being applied.

Beyond data protection and non-discrimination, providers and deployers show limited awareness of the fundamental rights risks that their systems may pose, in particular to rights that may be of relevance in specific high-risk areas. For example, none of the respondents in education mentions the right to education (Article 14 of the Charter), none of those in employment mentions the freedom to choose an occupation and the right to engage in work (Article 15 of the Charter) and none of those in law enforcement mentions the presumption of innocence and the right to defence (Article 48 of the Charter).

Some providers and deployers indicate that they need structured guidance, for example in the form of templates, to help them consider relevant questions and identify fundamental rights risks. Some respondents also stress the need for training to enable them to carry out fundamental rights risk assessments of AI and – in addition – point to the need for more guidance on bias detection. Such guidance can help to realise the AI Act's potential for ensuring responsible innovation, fundamental rights protection and a level playing field between providers in practice.

This means that the guidance accompanying the standards under Article 9 of the AI Act and the FRIA template under Article 27 of the AI Act needs to be sufficiently clear for the effective assessment of fundamental rights risks in practice.

Investment in studies and testing of AI systems, particularly in high-risk areas, will allow for a better understanding of fundamental rights risks and effective mitigation practices.

Provide an evidence base for assessing and mitigating fundamental rights risks

FRA OPINION 6

The Commission and Member States should invest in establishing an evidence base that allows for a better understanding of fundamental rights risks and effective mitigation practices. As part of the current financial investment in AI in the EU, research and testing facilities on the safe use of AI are essential. Investments need to be made in studies on and the testing of AI systems' compliance with fundamental rights, particularly in high-risk areas. Most notably, bias testing needs to be further developed to allow for a better understanding of when and how AI can be used in practice without infringing fundamental rights.

The research indicates that currently the understanding of how the development and use of AI systems can have an impact on fundamental rights in practice is limited. In addition, the approach to mitigating the fundamental rights risks that AI systems can pose, which is a key component of FRIAs, is fragmented.

Respondents have different views on what constitute mitigation measures, ranging from data protection and governance measures to bias testing, human oversight measures and external audits. While all of these can contribute to mitigating fundamental rights risks, there is a general absence of examples and evidence of effective mitigation measures for providers and deployers to draw from. Therefore, providers and deployers may not have a structured or informed approach to mitigating risks to fundamental rights.

The mitigation measures mentioned by respondents do not systematically address or fully cover the potential fundamental rights risks identified. The focus, in most cases, is on measures addressing specific risks to a very small number of fundamental rights, such as risks to data protection and privacy and risks of bias and discrimination.

The fieldwork results show a strong reliance on human oversight as a mitigation measure. The effectiveness of human oversight as a mitigation measure depends on how it is designed and applied (see Article 14 and Article 26(2) of the AI Act) and whether tendencies to over-rely on the outputs of an AI system (automation bias) are duly considered when factoring in what human oversight involves. Human oversight of AI outputs cannot be the sole (and is not necessarily the most effective) solution to rights compliance. This is particularly true in the absence of knowledge about how AI works in practice and the extent to which both AI systems and the humans overseeing them can be biased, alongside other fundamental rights considerations.

The report clearly highlights the limitations of self-assessments. Despite providers and deployers being more aware of some of the fundamental rights risks involved in using AI, they struggle to understand and assess risks. This is partly because many providers and deployers work on innovative technologies in high-risk areas, for which fewer precedents exist. This also means that they face a lack of knowledge and experience of how these systems work in practice.

Self-assessments matter, but they are effective only alongside independent oversight by adequately resourced bodies with expertise in fundamental rights.

Ensure proper oversight for effective fundamental rights protection in the context of artificial intelligence

There is limited awareness about the fundamental rights that are potentially impacted by the development and use of AI systems, as research interviews highlight. FRA findings also demonstrate the limits of self-assessments in terms of both the classification of AI systems under the AI Act and providers' and deployers' assessments concerning the fundamental rights risks of their systems.

Providers of the high-risk AI systems set out in Annex III to the AI Act are mostly required, pursuant to Article 43(1) and (2) of the AI Act, to follow a conformity assessment procedure based on internal control to demonstrate compliance with Article 9, including concerning their management of fundamental rights risks. Deployers can, under Article 27 of the AI Act, conduct the fundamental rights assessment on their own.

Notably, while the AI Act introduces a novel EU database for high-risk AI systems (Article 71 of the AI Act), which increases transparency, the results of these self-assessments do not necessarily need to be made (fully) publicly available (see Article 49(4) and Article 71 of and Annex VIII to the AI Act). This limits external scrutiny and points to the need for effective oversight to ensure a harmonised and fundamental-rights-compliant implementation of the AI Act. In this regard, some interviewees and focus group participants highlight the importance of external oversight, though this was not explicitly asked about in the fieldwork.

Existing human rights protection structures apply equally to the use of AI. In previous reports, FRA has called for existing human rights structures – including data protection authorities, equality bodies, national human rights institutions, ombuds institutions and consumer protection bodies – to be built on to address fundamental rights risks stemming from the use of AI. The AI Act assigns some of these bodies additional responsibilities. Notably, data protection authorities act as market surveillance authorities for high-risk AI systems in the areas of law enforcement, migration, asylum and border management, and justice and democracy (pursuant to Article 74(8) of the AI Act), including certain biometric systems in these areas. They also have reporting duties in relation to the use of 'real-time' remote biometric identification systems in publicly accessible spaces for law enforcement purposes under Article 5(6) of the AI Act.

In addition, Article 77 of the AI Act empowers public bodies responsible for protecting fundamental rights to access documentation created or maintained under the AI Act whenever it is necessary in order to effectively fulfil their mandates. In certain situations, these public bodies can also ask market surveillance authorities to organise technical testing of high-risk AI systems. Member States have provided an initial list of bodies that fall under this definition.

FRA OPINION 7

While self-assessment is important, it functions appropriately only in combination with effective oversight by independent bodies that are sufficiently resourced and possess the necessary fundamental rights expertise. The EU and its Member States should ensure sufficient financial, human and technical resources for bodies involved in the protection of fundamental rights, including those appointed under Article 77 of the AI Act. Particular attention should be paid to those bodies that will be taking up new tasks under the AI Act.

FRA called, in its **2020 report *Getting the Future Right – Artificial intelligence and fundamental rights***, on the EU and its Member States to ensure that these bodies receive sufficient resources, powers and expertise to oversee AI development and use. **FRA's 2024 report *GDPR in Practice – Experiences of data protection authorities*** shows that some data protection authorities are facing challenges relating to the resources needed to carry out their duties, including as a result of new duties assigned to them under EU law, such as the AI Act. Given the absence of knowledge among the interviewees about the full range of rights that can be impacted by AI, it is apparent that there will be increased demands on expert oversight bodies to provide advice on and to oversee high-risk AI systems with respect to fundamental rights compliance. These are tasks for which they may not have the requisite resources.

Introduction

Most people in the EU, namely 83 % according to a recent Eurobarometer survey, consider it important that public authorities shape the development of artificial intelligence (AI) and other digital technologies to ensure that they respect our rights and values ([1]). In 2025, the European Commission announced the largest public investment in AI in the world ([2]). The increase in the adoption of AI presents unprecedented opportunities and challenges for societies. AI systems, particularly those that influence important domains such as asylum, education, employment, law enforcement and public benefits, carry profound implications for fundamental rights ([3]). As these technologies increasingly influence decision-making, their alignment with the EU's fundamental rights framework needs to be ensured.

The EU institutions have recognised the risks AI poses to fundamental rights and undertaken a range of activities to address these risks. These efforts have led, among other initiatives, to the adoption of the AI Act ([4]). One of the objectives of this directly applicable regulation is to protect the fundamental rights enshrined in the Charter of Fundamental Rights of the European Union (the Charter) when developing and using AI in the EU (see recitals 1, 5, 6 and 7 and Article 1(1) of the AI Act).

The AI Act takes a risk-based approach to regulating AI ([5]). That means that different areas of application of AI face varying levels of regulation and requirements. In Article 5, it prohibits certain uses of AI, listing AI uses that are considered unacceptable in the context of EU values, such as manipulative or exploitative AI practices or social scoring. However, the bulk of the AI Act's provisions deal with high-risk AI systems, as defined in Article 6 of the AI Act. These high-risk systems are (1) products (or safety components of products) covered by the EU product safety legislation listed in Annex I to the

AI Act and requiring a third-party conformity assessment and (2) selected use cases in the areas of:

— biometrics;
— critical infrastructure;
— education and vocational training;
— employment, workers' management and access to self-employment;
— access to and enjoyment of essential private services and essential public services and benefits;
— law enforcement;
— migration, asylum and border management;
— administration of justice and democratic processes.

The AI Act includes a series of provisions that require providers of high-risk AI systems – those who develop systems and place them on the market – to help safeguard certain fundamental rights. Moreover, deployers of high-risk AI systems – those who use AI systems under their authority – are also subjected to certain requirements. These are further explained in Section **1.2**.

The AI Act is one piece of EU legislation that aims to support fundamental rights protection when AI is being developed and used. However, it is by no means the only relevant law. Besides any sectoral legislation that may apply, such as legislation regulating medical devices ([6]) and digital services ([7]), EU data protection law includes provisions that are relevant for AI, as the use of AI often involves processing personal data ([8]). Another example is the EU's non-discrimination legislation ([9]), which also applies to the use of AI.

This report focuses on the AI Act and how to implement it to best safeguard fundamental rights. It does not examine in detail the interplay of the AI Act with other applicable law, which might warrant a separate study and on which the Commission will provide guidance in the future ([10]).

In parallel to the EU's AI Act, in September 2024, the EU signed the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law ([11]). The framework convention outlines general principles related to activities within the AI life cycle and includes a provision on the assessment and mitigation of risks and adverse impacts (Article 16). In November 2024, the Council of Europe Committee on Artificial Intelligence adopted a non-binding methodology for the risk and impact assessment of AI systems from the point of view of human rights, democracy and the rule of law ([12]).

This report is based on 38 semi-structured interviews with potential providers and/or deployers of high-risk AI systems, and also experts in selected high-risk AI areas (asylum, education, employment, law enforcement and public benefits). The interviews focused on use cases appearing in one or more of the following five EU Member States: Germany, Ireland, the Netherlands, Spain and Sweden. These were selected to ensure a diverse sample of Member States with respect to the AI policy situation and the level of adoption of AI in the areas covered, and for the methodological feasibility of carrying out the research.

In addition, the report features insights from focus groups and interviews with rights holders, which focused on use cases in the areas of education, employment and law enforcement. These rights holders comprise 18 members of the public. Given the small number of people participating, the views expressed do not claim to be representative of rights holders' opinions, but rather serve to add further insights. The Annex to this report contains a description of the methodology.

The fieldwork addressed the following questions.

- When is a system considered a (high-risk) AI system by potential providers and deployers of AI in certain areas?
- What kinds of AI systems are being developed or are in use today and what fundamental rights risks do they carry? How are AI systems currently assessed and addressed in practice?
- What guidance do providers and deployers need in order to be able to effectively assess the fundamental rights impacts of their systems? What are the main elements of effective assessments?

By providing answers to these questions, this report aims to offer an empirical basis for the development of much-needed practical guidance. It builds on existing European Union Agency for Fundamental Rights (FRA) research in the field ([13]), and is further complemented by other FRA reports on the digitalisation of justice ([14]) and on the use of remote biometric identification systems for law enforcement purposes (forthcoming).

This report addresses developments, focusing on fundamental rights challenges, with respect to high-risk AI. **Chapter 1** discusses the definitions of an AI system and a high-risk AI system and how respondents classify their own systems in practice. **Chapter 2** describes the AI Act's requirements for assessing the fundamental rights risks of high-risk AI systems, current practices in assessing high-risk AI, fundamental rights risks in different high-risk areas and the mitigation measures taken. **Chapter 3** addresses specific challenges and opportunities involved in assessing high-risk AI, the guidance needed and the main elements of any form of effective fundamental rights assessment of high-risk AI. The concluding chapter reflects on the fieldwork findings and future steps needed for realising the AI Act's potential in ensuring responsible innovation, fundamental rights protection and a level playing field between providers in practice.

Endnotes

(1) European Commission, ***Special Eurobarometer 566 – The Digital Decade 2025***, Eurobarometer report, 2025, pp. 37 and 45–46.
(2) European Commission, '**Speech by President von der Leyen at the European working meeting of the Artificial Intelligence Action Summit**', European Commission website, 10 February 2025, accessed 8 October 2025.
(3) See, for example, European Union Agency for Fundamental Rights (FRA), ***Bias in Algorithms – Artificial intelligence and discrimination***, Publications Office of the European Union, Luxembourg, 2022; FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020; Rodrigues, R., '**Legal and human rights issues of AI: Gaps, challenges and vulnerabilities**', *Journal of Responsible Technology*, Vol. 4, 2020, 100005.
(4) **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)** (OJ L, 2024/1689, 12.7.2024.
(5) Applied AI, ***AI Act: Risk classification of AI systems from a practical perspective – A study to identify uncertainties of AI users based on the risk classification of more than 100 AI systems in enterprise functions***, Munich, 2023. Based on a review of over 100 systems used by enterprises, this report concludes that 1 % would be banned, 18 % would be considered high risk and 42 % would be considered low risk. For 40 %, it is unclear whether they would be considered high or low risk (pp. 12–13; it should be noted that these percentages amount to 101 %, which is presumably due to rounding). The report was published before the agreement on the final text of the AI Act, notably before the addition of the high-risk filter.
(6) **Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC** (OJ L 117, 5.5.2017, p. 1; **Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on *in vitro* diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU** (OJ L 117, 5.5.2017, p. 176.
(7) **Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act)** (OJ L 277, 27.10.2022, p. 1.
(8) This includes the General Data Protection Regulation (**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** (OJ L 119, 4.5.2016, p. 1), the Law Enforcement Directive (**Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA** (OJ L 119, 4.5.2016, p. 89) and the Data Protection Regulation for EU institutions, bodies, offices and agencies (**Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC** (OJ L 295, 21.11.2018, p. 39).
(9) **Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin** (OJ L 180, 19.7.2000, p. 22; **Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation** (OJ L 303, 2.12.2000, p. 16; **Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services** (OJ L 373, 21.12.2004, p. 3; **Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast)** (OJ L 204, 26.7.2006, p. 23. See also FRA, European Court of Human Rights and Council of Europe, ***Handbook on European Non-discrimination Law***, Publications Office of the European Union, Luxembourg, 2018.
(10) Article 96(1)(e) of the AI Act.
(11) Council of Europe, **Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law**, Council of Europe Treaty Series, No 225, 2024 (adopted 17 May 2024, opened for signature 5 September 2024).
(12) Council of Europe Committee on Artificial Intelligence, '**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)**', CAI(2024)16rev2, Strasbourg, 28 November 2024. This methodology will be complemented by a model, which will contain supporting materials. See Council of Europe, '**Huderia – Rrisk and impact assessment of AI systems**, Council of Europe website, accessed 8 October 2025.
(13) FRA, ***#BigData: Discrimination in data-supported decision making***', Publications Office of the European Union, Luxembourg, 2018; FRA, ***Data Quality and Artificial Intelligence – Mitigating bias and error to protect fundamental rights***, Publications Office of the European Union, Luxembourg, 2019; FRA, ***Facial Recognition Technology: Fundamental rights considerations in the context of law enforcement***, Publications Office of the European Union, Luxembourg, 2019; FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020; FRA, ***Bias in Algorithms – Artificial intelligence and discrimination***, Publications Office of the European Union, Luxembourg, 2022.
(14) FRA, ***Digitalising Justice: A fundamental-rights-based approach***, Publications Office of the European Union, Luxembourg, 2025.

1

ARTIFICIAL INTELLIGENCE SYSTEMS AND THEIR CLASSIFICATION AS HIGH RISK

This chapter describes the AI Act's definition of AI and the requirements for the classification of AI systems as high risk. It then outlines the high-risk areas and use cases covered in this report. Finally, the chapter describes how respondents view their own systems and whether they would classify them as AI and as potentially high risk.

1.1. DEFINITION OF ARTIFICIAL INTELLIGENCE

The definition of AI is crucial to the AI Act, as this determines the material scope of the regulation (i.e. whether it is applicable to a particular case). FRA highlighted in its 2020 report on AI and fundamental rights that developers and users of AI have very different understandings of what AI is or should be. While some have a broader understanding of AI, others consider only certain technologies to be AI. In 2020, FRA called for a clear definition of AI to provide legal clarity, but also for a regular review of the definition. As AI can be understood in different ways, the technologies usually covered vary considerably with respect to their complexity (i.e. the number and combination of parameters used) and level of automation (i.e. the degree of human review of outputs). Furthermore, the scale of application, and also the potential impact and harm, vary across different technologies ([15]). FRA highlighted in 2023 that discussions around the definition of AI should avoid narrowing its scope, as this may unduly narrow the AI Act's scope of protection ([16]).

The term 'AI system' is defined in Article 3(1) of the AI Act. However, despite there being a definition and guidance in place, challenges remain in interpreting the scope of the definition, as further explained below.

'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

Article 3(1) of the AI Act

The definition of an 'AI system' in the AI Act is based on the Organisation for Economic Co-operation and Development's updated definition of an AI system. It covers various elements, including (1) machine-based systems, (2) autonomy, (3) adaptiveness, (4) AI system objectives, (5) inference how to generate outputs, (6) such as predictions, content, recommendations, or decisions, and (7) interaction with the environment. The definition still leaves much room for interpretation. For example, 'designed to operate with varying levels of autonomy' may mean that there is a minimum level of autonomy, potentially leaving some systems outside its scope ([17]), as also indicated by recital 12 of the AI Act. The ability to 'infer' is one of the key features of AI systems. It is not defined in the provisions of the AI Act, but recital 12 clarifies that inference 'refers to the process of obtaining the outputs … and to a capability of AI systems to derive models or algorithms, or both, from inputs or data' ([18]). Recital 12 of the AI Act also states that the capacity to infer goes beyond basic data processing and provides some examples (a non-exhaustive list) of techniques that enable inference (e.g. machine learning approaches).

Some of the individuals interviewed for this report highlight challenges in relation to the definition of AI in the AI Act, including that it does not provide enough clarity. A deployer in the area of public benefits and one expert on AI and employment consider the definition to be quite broad. The latter explains this as follows: 'basically any algorithm that leads to an actual change or generates content is considered an AI system'. This interviewee explains that AI is closely related to analytics, as analytics is used to draw conclusions: 'Even if you're just analysing something, it could already be considered an AI system, as it generates content. Especially when an algorithm follows data analysis and leads to actual changes.'

Another deployer in the area of asylum considers the definition to not be sufficiently precise. They consider that the definition would, in their view, ultimately amount to software projects that use a model that has been trained on data.

A respondent who is an expert on AI and law enforcement does not consider the system that they helped to develop to fall under the AI Act, as it lacks autonomy. Another respondent from a provider in the area of employment does not consider their system to be an AI system under the AI Act, as it does not include deep learning. This respondent, however, mentions that a narrow understanding of the definition (as they see it) is not useful, because even their simple system, which may not be covered by the AI Act, poses threats to fundamental rights.

Another provider in the area of employment mentions that the definition may push people to use simpler systems that would not classify as AI, even though these systems can have similar fundamental rights risks:

> *I think algorithms that are not AI can be just as harmful, and so it is pushing people to then say: oh well, we do not do AI, we actually just have, you know, regression tools. However, these regression tools can be just as problematic.*
>
> Provider of an AI system in the area of employment

Issues concerning the definition were also raised during the European Commission's public consultation on the definition of AI systems. While

many respondents considered that traditional statistical models do not meet the criteria contained in Article 3(1) of the AI Act, several respondents 'raised concerns about potential loopholes in the AI Act that could allow developers to circumvent regulations by reclassifying AI systems as traditional software' ([19]).

During the public consultation, the European Law Institute, for example, discussed the definition's vagueness and lack of clarity ([20]). Civil-society organisations called for a confirmation that comparatively 'simple' systems are within the scope of the definition, as the regulation has to focus on potential harms and not just technical methods ([21]).

At the beginning of 2025, after the fieldwork for this report had been completed, the European Commission published guidelines on the definition of an AI system established in the AI Act ([22]). In the guidelines, an effort is made to narrow the definition. For example, systems for improving mathematical optimisation are presented as outside the scope of the definition of AI. This includes established methods, such as linear and logistic regression, that have been used for many years to improve the efficiency of optimisation algorithms in computational problems. However, this description remains open to interpretation, particularly because logistic regression is a commonly used machine learning algorithm.

The guidelines on the definition have been subject to criticism and requests for further clarification ([23]). One of the issues raised is the potential harm posed by 'simpler' systems that are excluded from the definition ([24]). Without further clarity, the Court of Justice of the European Union (CJEU) may need to clarify the definition of AI and the scope of the AI Act. It is clear that, based on the interviews and analysis conducted for this report and the statements made by other stakeholders, the definition of AI remains open to interpretation, as it can be understood in either a broader or a narrower way by different actors.

1.2. HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEMS UNDER THE ARTIFICIAL INTELLIGENCE ACT

The AI Act defines high-risk AI systems as systems that either:

— are products (or safety components of products) covered by the EU product safety legislation listed in Annex I to the AI Act and requiring a third-party conformity assessment;
— fall under the use cases listed in Annex III to the AI Act ([25]).

Annex III to the AI Act lists eight high-risk areas of application, each of them detailing specific purposes of use. These are:

— biometrics;

— critical infrastructure;
— education and vocational training;
— employment, workers' management and access to self-employment;
— access to and enjoyment of essential private services and essential public services and benefits;
— law enforcement;
— migration, asylum and border control management;
— administration of justice and democratic processes.

This means that AI systems used for specific purposes in those areas will generally be considered high-risk AI systems. Providers and deployers of such systems will need to comply with the requirements contained in Chapter III of the AI Act.

Table 1 sets out some of the requirements of the AI Act that can contribute to safeguarding fundamental rights. The risk management system for providers and the fundamental rights impact assessments (FRIAs) that have to be undertaken by certain deployers will be explained further in Section **1.4**.

TABLE 1: SELECTED REQUIREMENTS FOR PROVIDERS AND DEPLOYERS OF HIGH-RISK AI SYSTEMS FOR SAFEGUARDING FUNDAMENTAL RIGHTS

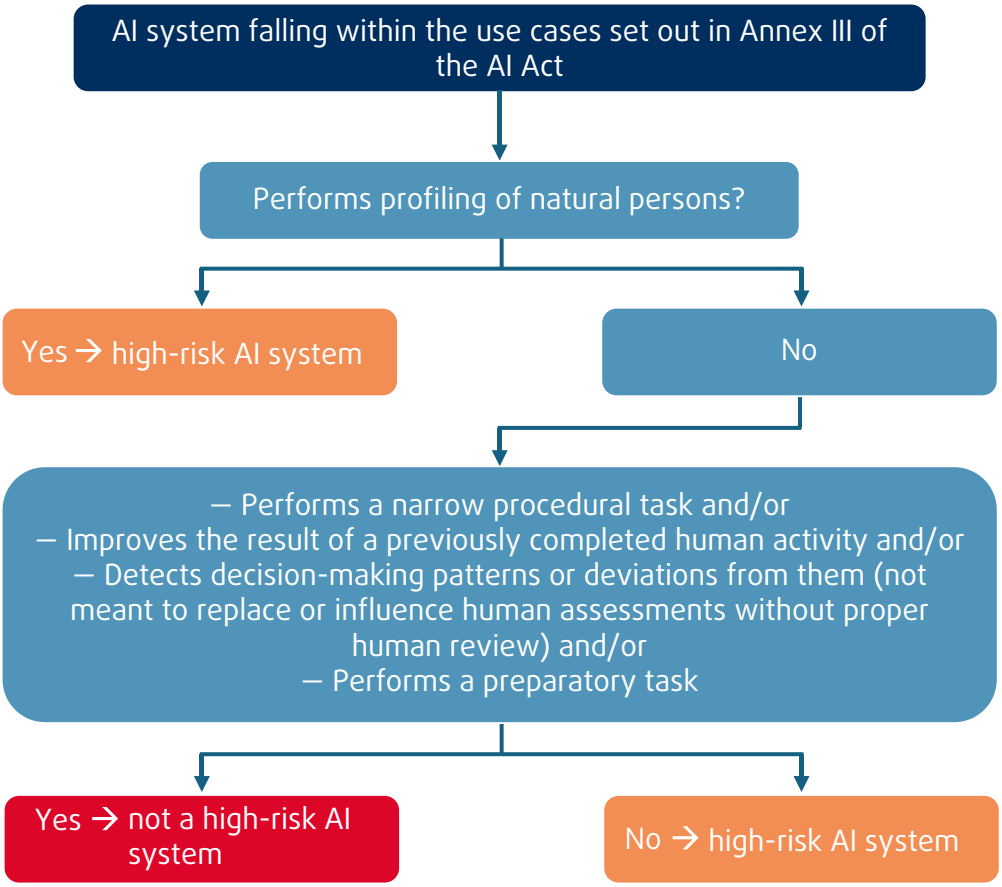
Requirement	Who	Key points
Risk management system (Article 9 of the AI Act)	Providers	• Needs to be in place throughout the entire life cycle of the system • Involves the identification and analysis of the known and reasonably foreseeable risks to health, safety and fundamental rights when the system is used according to its intended purpose and under conditions of reasonably foreseeable misuse • Involves the adoption of appropriate and targeted risk management measures
Data governance (Article 10 of the AI Act)	Providers	• Involves the examination of training, validation and testing datasets for possible biases that may have an impact on people's health, safety or fundamental rights • Involves the identification of appropriate measures for the detection, prevention and mitigation of such biases • Involves the processing of special categories of personal data relating to protected characteristics when strictly necessary for the identification and correction of biases, subject to strict conditions and safeguards
Technical documentation (Article 11 of and Annex IV to the AI Act)	Providers	• Needs to be drawn up before the system is placed on the market / put into service and then kept up to date • Needs to demonstrate compliance with the requirements for high-risk AI systems
Transparency and provision of information to deployers (Article 13 of the AI Act)	Providers	• Providers need to provide deployers with detailed and comprehensible instructions for use, including concerning the characteristics, capabilities and limitations of the system, its performance regarding particular groups of people (when appropriate) and information to enable the interpretation of the system's output (where applicable)
Human oversight (Articles 14 and 26 of the AI Act)	Providers and deployers	• Providers need to develop systems in a way that allows for human oversight while they are in use • Deployers need to assign the responsibility for human oversight to people with the necessary competencies, training, authority and support
FRIA (Article 27 of the AI Act)	Certain deployers	• Needs to be conducted prior to the deployment of the system and updated when needed • Needs to cover various elements, including the people likely to be affected, specific risks of harm, human oversight measures and the mitigation measures planned if the risks materialise

Source: FRA analysis of the AI Act, 2025.

Several requirements for high-risk AI systems in the AI Act will be further explained through guidance provided by the European Commission ([26]). In addition, for certain requirements for providers, the European Commission requested that the European Committee for Standardization (CEN) and the European Committee for Electrotechnical Standardization (CENELEC) develop technical standards (see Section **1.4.1**) ([27]).

When it comes to the classification of AI systems as high risk, Article 6(3) of the AI Act contains a derogation (also called a '**filter**' and referred to as such in the report), whereby the AI systems specified in Annex III to the AI Act 'shall not be considered to be high-risk where [they do] not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making'. This applies if one or more of the four conditions in Article 6(3) of the AI Act are fulfilled, as shown in **Figure 1**. Nevertheless, an AI system is considered high risk if it undertakes the profiling of natural persons (i.e. is used to evaluate certain personal aspects of a person based on their personal data, for example to predict their behaviour or health), which is defined in accordance with the General Data Protection Regulation (GDPR) ([28]).

FIGURE 1: FILTER APPLYING TO THE CLASSIFICATION OF AI SYSTEMS AS HIGH RISK



Source: FRA visualisation of Article 6(3) of the AI Act

Providers of the AI systems listed in Annex III to the AI Act who consider that their systems are **not** high risk (i.e. the systems do not perform profiling **and** one or more of the four conditions set out under the filter are fulfilled) must document their assessment and register themselves and their systems in the upcoming EU database. This database will contain information about the high-risk AI systems listed in Annex III and about the AI systems to which the filter was applied ([29]).

Depending on how these criteria are expanded upon by the Commission and interpreted by the CJEU, and also how they are implemented in practice, the filter clause may limit the number of high-risk AI systems covered by the AI Act. The Commission is empowered by Article 6(6) and (7) of the AI Act to amend the filter's criteria through delegated acts. The scope of the definition of high-risk AI systems may therefore change in the future, if the Commission decides to exercise this prerogative.

At the time of finalising the fieldwork for this report, there was little guidance on what an assessment of whether a system is high risk or not may look like in practice. The Commission, in consultation with the AI Board, will have to provide guidelines on the application of Article 6 by 2 February 2026 ([30]). These guidelines will also be informed by a public consultation, which the Commission launched in June 2025 ([31]).

To conclude, for a system to be considered a high-risk AI system under the AI Act, the system will (1) have to fulfil the criteria of the definition of an 'AI system' (see section **1.1**), (2) fall under Annex I or Annex III to the AI Act and (3) and in case of Annex III systems pose a significant risk of harm to people's health, safety or fundamental rights. Further material and temporal exceptions may apply. For example, the AI Act does not apply to AI systems used exclusively for military, defence or national security purposes or to AI systems or models that are specifically developed and put into service solely for scientific research and development purposes ([32]). The AI Act will apply to high-risk AI systems already on the market only if they are subject to a significant change in their design. If public authorities use such systems, system compliance with the AI Act must be ensured by 2 August 2030 ([33]).

Importantly, the specific uses of some AI systems might result in them being considered either high risk or prohibited under the AI Act. For example, while emotion recognition systems used in workplaces and education institutions are generally prohibited, their use in other areas is considered high risk, for example in the context of migration or law enforcement ([34]). Some AI practices may also be prohibited under other applicable EU law ([35]).

1.3. ARTIFICIAL INTELLIGENCE USE CASES COVERED IN THIS REPORT

The present report covers different AI use cases in the areas of asylum, education, employment, law enforcement and public benefits. The main use cases examined are highlighted in **Table 2**. Many of these would likely be considered as high-risk AI use cases under the AI Act, subject to the definitions and derogation described earlier. It should be noted that further use cases, not featured in **Table 2**, were discussed with interviewees. Insights from these complement the report's findings.

TABLE 2: OVERVIEW OF THE AI USE CASES INVESTIGATED IN THIS REPORT

Area	Use case	Possible link to Annex III to the AI Act
Asylum	AI language assessment for origin determination in asylum procedures	Point 7(c): if used for the examination of the eligibility for asylum, including the reliability of evidence Point 1(b): if used for certain types of biometric categorisation
	AI system supporting the search for country-of-origin information	Point 7(c): if used for the examination of the eligibility for asylum, including the reliability of evidence Point 8(a): if part of a judicial system to research and interpret facts and the law and to apply the law to a concrete set of facts
	AI system for detecting security-related information in asylum hearings that should be transmitted to security authorities	Point 7(b): if used to assess a security risk posed by a person who entered into the territory of a Member State Point 6: if used for certain law enforcement purposes
Education	AI system for assessing reading ability in early education	Point 3(a): if used to determine access, admission or assignment to educational and vocational training institutions Point 3(b): if used to evaluate learning outcomes, including if those outcomes are used to steer learning processes Point 3(c): if used to assess the appropriate level of education that a person will receive / be able to access
Employment	AI-based search engine for recommending, ranking or shortlisting job applicants	Point 4(a): systems used for the recruitment or selection of people, in particular for analysing and filtering job applications
	AI system for supporting hiring decisions by assessing the cognitive, social and emotional competencies of candidates	Point 4(a): systems used for the recruitment or selection of people, in particular for evaluating candidates
	AI system for recommending jobs to jobseekers based on their backgrounds, preferences and other input	Point 4(a): systems used for the recruitment or selection of people
Law enforcement	AI system for assessing the likelihood of complaints to the police being false, mostly in relation to insurance fraud in the context of reports of theft and burglaries	Point 6(a): if used to assess the risk of a person becoming a victim of crime Point 6(c): if used to evaluate the reliability of evidence in the course of the investigation or prosecution of criminal offences
	AI system for assessing a range of risks related to prisoners' behaviour, from the risk of recidivism to the risk of violence towards other prisoners and self-harm	Point 6(a): if used to assess the risk of a person becoming a victim of crime Point 6(d): if used to assess the risk of a person offending or re-offending not solely based on profiling or to assess personality traits and characteristics or past criminal behaviour
Public benefits	AI system for supporting the assessment of social insurance applications by determining health-related classifications based on patients' health data	Point 5(a): if used to evaluate the eligibility for or to grant, reduce, revoke or reclaim essential public assistance benefits and services
	AI system for transcribing and summarising conversations between applicants for public benefits and case workers in public administration	Point 5(a): if used to evaluate the eligibility for or to grant, reduce, revoke or reclaim essential public assistance benefits and services

Note: To guarantee the anonymity of respondents, no references to specific AI systems and details about how they function are included.

Source: Desk research and interviews undertaken for this study.

1.4. CLASSIFICATION OF HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEMS IN PRACTICE

During the fieldwork for this report, the AI Act did not yet apply to high-risk systems and further guidance on its application was not yet available from the AI Office. This is also reflected in the interviewees' responses, which indicate diverging views and uncertainties about the high-risk classification in practice. A significant number of interviewees are unsure about the classification or consider that it would depend on the specific system's functionalities and context of use and the interpretation of the AI Act's provisions.

1.4.1. Interviewees' views on the classification

The initial mapping for the research suggested that the systems explored could be high risk, which is not always confirmed by the interviewees. Some interviewees ultimately agree that their systems would be high risk, while others do not think that they would be and yet others cannot decide. In some cases, respondents, when referring to the same system, have different views on the classification. This highlights the different interpretations of what constitutes a high-risk AI system, with other possible reasons for this mismatch including interviewees' background knowledge about the system and about the AI Act. In addition, some respondents consider that the system discussed would not amount to an AI system under the AI Act, for example because it does not fulfil the criteria for an 'AI system' under Article 3(1) of the AI Act (see Section **1.3**).

In a few instances, respondents working for both providers and deployers give reasons that are not linked to the AI Act definition for their conclusion that the AI system discussed is not high risk. For example, one interviewee does not consider their system to be high risk because it does not process personal data. Another mentions that their system is not high risk because any risks are ruled out through the various mitigation measures in place. One respondent working for the provider of a system used to identify security-related information in asylum hearings points out that some confusion arises from the fact that Article 2(3) of the AI Act contains an exception for national security, but that other provisions seem to apply to security-related systems.

When asked about the high-risk classification and the filter under Article 6(3) of the AI Act, interviewees are sometimes unsure if and how the classification applies. Some interviewees consider that their system would fall under the filter. For example, a provider of a system for assessing reading ability in early education considers that their system performs a narrow procedural task, being a small part of a larger system. A deployer also says that their AI system should not be considered high risk, as it fulfils a narrow procedural task that is part of a wider assessment process.

Three experts express concerns about the filter being interpreted too broadly. One of them, with respect to the migration and asylum area, points out that the phrasing in Article 6 makes it seem that systems that do not influence the final decision are not high risk. This expert emphasises that, even in such cases, a system could still pose significant risks to asylum seekers by rendering them non-credible in cases of errors in the system.

It is therefore necessary to look at the various use cases of a system to assess what kind of impact it could have before categorising it as (non-) high risk. Two other experts highlight that the uncertainty surrounding the classification can lead to different classifications in practice, depending on the company culture, or to resistance and reluctance to use high-risk AI systems.

Anything where you're talking about education and children, it's always going to be high risk. And I think that special attention needs to be put into these systems, in the building of these systems to make sure that we don't allow bad technology [to] infiltrate the school system.

Provider of an AI system in the area of education

Some interviewees emphasise the inherent high-risk nature of systems used in certain areas. A respondent working for a provider and deployer of an AI system for recommending jobs to jobseekers considers that their system is high risk, as it falls under the definition of AI and can have an impact on people's work and lives. They emphasise the inherent high-risk nature of all AI systems operating within the recruitment and selection domain.

A respondent working for a provider and deployer of an AI system for supporting the assessment of social insurance applications explains that the decision-making process of determining health-related classifications is inherently high risk. They note that this is regardless of whether AI is involved or not, as incorrect decisions can have a substantial effect on a person's life, such as potentially not granting social insurance to people who have difficulties working.

[The system determining health-related classifications] is essentially a decision support system, so it is an individual who makes the decision. On the other hand, one could argue that perhaps one just presses the button and accepts what ... the AI suggests, and then it becomes, in a sense, an AI decision. On the other hand, it might be more of an educational issue ... that should be addressed in a different way. And here it becomes a bit uncertain whether it affects the classification: is this high risk or is it something else?

Provider and deployer of an AI system in the area of public benefits

Some interviewees suggest that the high level of human involvement in the decision-making process is a reason not to consider a system high risk. The filter determines that an AI system is not high risk if it does not materially influence the outcome of decision-making and lists four conditions where this is the case (see earlier). Some interviewees consider that their systems may not be high risk, as the system does not, among other things, automate final decisions. People who are responsible for the decision can override the outcome.

However, some interviewees observe that there is a danger of automation bias, with one provider in public benefits pointing out that human-made decisions can also be faulty and biased. On the other hand, two respondents from a provider and deployer developing a broader system aimed at handling the administrative work of public benefits case managers, including through transcription and summary tools, consider that the system they are developing is high risk, as it will contribute to decision-making processes. The nuances in how high-risk applications of AI are implemented (i.e. whether and to what extent they are used for supporting or automating decisions) leaves room for interpretation about whether the system can be classified as high risk.

1.4.2. Limits of self-assessments

In practice, providers whose system would generally fall under Annex III to the AI Act will have to self-assess whether it poses a significant risk of harm to people's fundamental rights and ultimately whether it is to be considered high risk. Article 6(4) of the AI Act sets out that providers need to document their assessment before they put such a system on the market or into service. They need to register the system in the EU database that will contain information about the high-risk AI systems listed in Annex III and about AI systems to which the filter was applied. This will include indicating which of the four conditions listed in Article 6(3) of the AI Act is fulfilled for the exception to apply. A short summary of the grounds for this conclusion will need to be given ([36]).

This documentation will be publicly accessible, except in the areas of law enforcement, migration, asylum and border control management. In these areas, the information will be accessible only to the Commission and designated national supervisory authorities ([37]). In these areas, providers will also not be required to register the applicable conditions from Article 6(3) of the AI Act or provide a summary of the grounds ([38]).

This self-assessment is against criteria that may leave room for interpretation and thus can introduce a high degree of subjectivity into the assessments. This may lead to misclassifications of systems and, therefore, may create fundamental rights risks that would otherwise be addressed by complying with provisions on high-risk AI systems. While providers will generally need to document the application of the filter, detecting misclassifications will depend on effective oversight in practice. The European Parliament's Legal Service ([39]), civil-society organisations ([40]) and representatives of academia ([41]) also raised concerns over the filter and providers' self-assessments.

The use of AI systems can generally have an impact on a wide spectrum of fundamental rights, regardless of the field in which the systems are applied ([42]). To this end, it should be noted that those AI systems that are self-assessed as falling under the filter conditions, and thereby not considered high risk, will still be used in high-risk areas that can have an impact on people's fundamental rights.

Endnotes

(15) FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020, pp. 26–28.
(16) FRA, ***Fundamental Rights Report – 2023***, Publications Office of the European Union, Luxembourg, 2023, p. 190.
(17) Bobev, T., '**Defining AI in the AI Act: Pin the tail on the system**', KU Leuven website, 2 April 2024, accessed 8 October 2025.
(18) This is also mirrored in Organisation for Economic Co-operation and Development (OECD), '**Explanatory memorandum on the updated OECD definition of an AI system**', *OECD Artificial Intelligence Papers*, No 8, March 2024, OECD Publishing, Paris, p. 9.
(19) Centre for European Policy Studies (CEPS), ***Analysis of EU AI Office Stakeholder Consultations: Defining AI systems and prohibited applications***, Publications Office of the European Union, Luxembourg, 2025, p. 23.
(20) European Law Institute (ELI), ***Commission guidelines on the application of the definition of an AI system and the prohibited AI practices established in the AI Act – Response of the European Law Institute***, Vienna, 2024, p. 7.
(21) AI Act Civil Society Coalition and #ProtectNotSurveil Coalition, '**Human rights and justice must be at the heart of the upcoming Commission guidelines on the AI Act implementation**', p. 1.
(22) **Communication from the Commission – Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)**, C(2025) 5053 final, of 29 July 2025.
(23) See Irish Council for Civil Liberties, '**AI system definition guidelines fail to provide clarity**', Irish Council for Civil Liberties website, 7 February 2025, accessed 8 October 2025; Wazir, R., '**The European Commission's guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act) – Comment by the Advisory Committee on the Ethics of Artificial Intelligence of the Austrian Commission for UNESCO**', Advisory Committee on the Ethics of AI, Austrian Commission for UNESCO, Vienna, 21 March 2025.
(24) Wazir, R., '**The European Commission's guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act) – Comment by the Advisory Committee on the Ethics of Artificial Intelligence of the Austrian Commission for UNESCO**', Advisory Committee on the Ethics of AI, Austrian Commission for UNESCO, Vienna, 21 March 2025, pp. 2–3.
(25) Article 6 of the AI Act.
(26) See, for example, Articles 6(5), 27(5) and 73(7), Article 96 and Article 112(1) of the AI Act.
(27) **Commission implementing decision of 22.5.2023 on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence**, C(2023) 3215 final of 22 May 2023.
(28) Article 3(52) of the AI Act and Article 4(4) of the GDPR. The Law Enforcement Directive and the Data Protection Regulation for EU institutions, bodies, offices and agencies include identical definitions of profiling to that in the GDPR. See Article 3(4) of the Law Enforcement Directive and Article 3(5) of the Data Protection Regulation for EU institutions, bodies, offices and agencies.
(29) Article 6(4) of the AI Act. See also Article 49(2) and Article 71 of the AI Act.
(30) Article 6(5) of the AI Act.
(31) European Commission, '**Commission launches public consultation on high-risk AI systems**', European Commission website, 6 June 2025, accessed 8 October 2025.
(32) Article 2(3) and (6) of the AI Act. See Article 2(4), (8), (10) and (12) of the AI Act for further material exceptions.
(33) Article 111(2) of the AI Act. For further rules on the temporal scope of the AI Act, see Article 111(1) and (3) of the AI Act.
(34) Article 5(1)(f) of and Annex III, point (1)(c), to the AI Act. For an explanation of the interplay between the prohibitions and the requirements for high-risk AI systems, see **Communication from the Commission – Commission guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)**, C(2025) 5052 final of 29 July 2025, Section 2.6.
(35) Article 5(8) of the AI Act. See also **Communication from the Commission – Commission guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)**, C(2025) 5052 final of 29 July 2025, paragraph 43.
(36) Articles 6(3) and 49(2) and Article 71 of and Annex VIII, Section B, to the AI Act.
(37) Articles 49(4) and 74(8) of the AI Act.
(38) Article 49(4)(b) of and Annex VIII to the AI Act.
(39) European Parliament's Legal Service, '**Non-paper – AI Act – Articles 6 and 7, and Annex III**', EP-PE_AVS(2023)3467, 13 October 2023, p. 3.
(40) Civil-society organisations, '**EU legislators must close dangerous loophole in AI Act**', joint statement, 7 September 2023, p. 1; European Digital Rights (EDRi) and AI coalition partners, '**EU's AI Act fails to set gold standard for human rights**', 3 April 2024, p. 6.
(41) See, for example, Wachter, S., '**Limitations and loopholes in the EU AI Act and AI liability directives: What this means for the European Union, the United States, and beyond**', *Yale Journal of Law & Technology*, Vol. 26, Issue 3, 2024, pp. 684–686.
(42) FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020, p. 7.

2

ASSESSING HIGH-RISK ARTIFICIAL INTELLIGENCE WITH RESPECT TO FUNDAMENTAL RIGHTS

This chapter outlines the main assessment requirements in relation to fundamental rights – for both providers and deployers – for high-risk AI systems under the AI Act. These requirements will largely apply as of 2 August 2026. The chapter then dives into how high-risk AI systems are currently being assessed in practice. Next, the chapter outlines the possible fundamental rights risks of use cases covered in the different high-risk areas. Finally, the chapter describes current mitigation practices.

Current assessment practices exhibit a strong focus on data protection, technical aspects and business or legal risks. Beyond data protection, other fundamental rights are normally not assessed in a structured manner. Compared with findings in FRA's 2020 *Getting the Future Right* report, providers and deployers do now seem to be more aware of potential fundamental rights risks beyond data protection risks, most notably risks of bias and non-discrimination.

At the same time, there may be many other potential risks to fundamental rights to consider, depending on the high-risk areas and use cases in question. Furthermore, the mitigation measures mentioned by respondents do not systematically address or fully cover the potential fundamental rights risks identified.

2.1. ARTIFICIAL INTELLIGENCE ACT REQUIREMENTS FOR ASSESSING FUNDAMENTAL RIGHTS

The AI Act includes provisions that oblige providers and certain deployers of high-risk AI systems to assess their systems' risks to fundamental rights, as laid down in the Charter. Providers and deployers have different sets of obligations in this regard under Articles 9 and 27 of the AI Act.

Providers' obligation to identify and analyse high-risk AI systems' risks to fundamental rights

Providers of high-risk AI systems are obliged by Article 9 of the AI Act to establish a risk management system that runs throughout the entire life cycle of the high-risk AI system. It involves several steps, among others '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'.

The risk management framework for providers is elaborated through standards developed by CEN and CENELEC, two European standardisation organisations. Compliance with the technical standards will result in a legal presumption that the provider's AI system conforms with the AI Act's requirements. While compliance with the standards will be voluntary, providers will be incentivised to follow them, as they will be able to self-assess their systems' conformity through an internal control measure.

At the same time, due to the AI Act's focus on product safety and fundamental rights, it is important to ensure that fundamental rights are well reflected and operationalised in these standards. Cooperation with fundamental rights organisations and standardisation bodies is crucial in this respect.

Sources: Articles 9, 40 and 43(1) of the AI Act; **Commission implementing decision of 22.5.2023 on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence**, C(2023) 3215 final of 22 May 2023; Smuha, N. A. and Yeung K., '**The European Union's AI Act: Beyond motherhood and apple pie?**', *SSRN*, 24 June 2024, pp. 27–28, 30 and 32. See also Gornet, M. and Maxwell, W., '**The European approach to regulating AI through technical standards**', *Internet Policy Review*, Vol. 13, Issue 3.

FRIAs conducted by certain deployers of high-risk AI systems

Certain deployers of high-risk AI systems listed in Annex III to the AI Act, except for those deploying systems in the area of critical infrastructure, are obliged by Article 27 of the AI Act to perform a FRIA. This obligation will apply to:

- bodies governed by public law;
- private entities providing public services;
- deployers of AI systems for evaluating creditworthiness or establishing a credit score, except for AI systems intended for financial fraud detection;
- deployers of AI systems used for risk assessment and pricing in life and health insurance.

These deployers will need to conduct the assessment prior to the deployment of the system and update it if there are any changes to the FRIA's elements. The assessment needs to cover certain elements pursuant to Article 27 of the AI Act, namely:

- a description of the processes in which the AI system will be used;
- the period of time and frequency of use;
- the categories of people and groups likely to be affected by its use;
- specific risks of harm;
- a description of the implementation of human oversight measures;
- the mitigation measures to be taken if the risks materialise.

The deployer can thereby rely on their data protection impact assessment, previously conducted FRIAs or existing impact assessments carried out by the provider. Article 13 of the AI Act serves as a link between the providers and deployers, as it obliges providers to submit certain information to deployers. This includes information about the performance of the high-risk AI system regarding specific people or groups of people, its level of accuracy and the conditions of use or foreseeable misuse that may pose risks to fundamental rights.

The deployer should notify the market surveillance authority of the results of the assessment in accordance with a template provided by the AI Office. FRA is assisting the AI Office in the development of the FRIA template.

A summary of the FRIA findings will also have to be registered in the publicly accessible EU database of high-risk AI systems. Exceptions to this exist for AI systems in the areas of law enforcement, migration, asylum and border control management.

Sources: Articles 13 and 27, Article 49(3) and (4) and Article 71 of the AI Act and Annex III to and Annex VIII, Section C(4), to the AI Act.

In the light of the requirements for a FRIA under Article 27 of the AI Act, it should be noted that not all deployers of the high-risk AI systems studied in the present report will have to conduct a FRIA. However, if a system falls into a high-risk area, all provider obligations apply. Significantly, a system falling outside the scope of the AI Act's definition of high-risk systems does not mean that the use of the system is without risks. Assessing fundamental rights when using AI and AI-related technologies is always good practice and leads to better-performing technology.

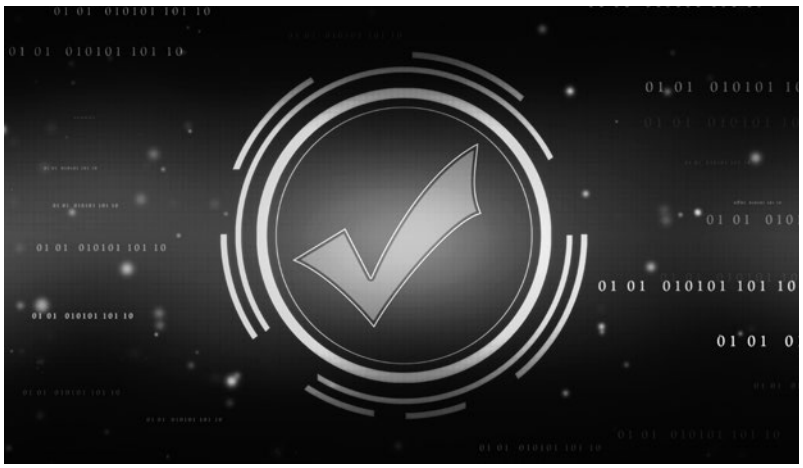
2.2. CURRENT PRACTICES IN ASSESSING (HIGH-RISK) ARTIFICIAL INTELLIGENCE

Across high-risk areas and use cases, organisations are generally not yet carrying out risk assessments that systematically take fundamental rights risks into account. There is a strong focus on data protection, technical aspects and business or legal risks. This is in line with earlier findings in FRA's 2020 *Getting the Future Right* report, in which a focus on data protection and technical aspects was reported in the context of AI impact assessments undertaken at the time.

Often the assessments are informal. For example, in the area of asylum, one of the deployers of an AI system that supports the search for country-of-origin information indicates that one of the 'unspoken premises' is that the rights to asylum, fair trial and non-discrimination are not compromised by the AI system. In addition, a provider in the area of asylum mentions that the impacts of the system on fundamental rights were discussed at an early stage and that the adoption of the AI Act has recently brought up the discussion again. A respondent from another provider mentions that they have quarterly meetings to discuss the assumptions that they made when building their test set:

> *So, we just did not formally call it a risk management system … but it operated as a risk management system. We just did not have a formal designation of that within the company, but it was something that we all cared about extensively.*
>
> Provider of an AI system in the area of employment

Two respondents from a provider and deployer of a system in the area of public benefits note that the AI Act's FRIA obligation is not yet clearly defined. They also note that (at the time of the fieldwork) there is a lack of a clear template explaining what they should do. One respondent from a provider states that they have not had enough interest from customers regarding fundamental rights issues to drive the development of internal guidelines within the organisation.



Sometimes a more structured approach to assessment is taken. In the area of education, a provider indicates that they analyse risks using a structured method, ensuring that they are aligned with both the GDPR and the AI Act. They add that criteria for assessing AI risks are constantly evolving, making it difficult to keep up.

Fundamental rights other than data protection are typically not considered. Current findings do show increased awareness of potential impacts on other rights, most notably non-discrimination. However, these are identified and assessed in an ad hoc and unstructured manner.

For example, one respondent from a deployer of an AI system in education mentions that they are currently developing a process to better detect and address potential discrimination in the AI systems they use, based on a checklist or set of questions. A respondent from a deployer in employment states that they undertook certain internal assessments to measure bias in terms of gender and age, and this increased awareness but unfortunately did not lead to the start of new processes, measures or policies. In addition, a deployer of a system that assesses a range of risks related to prisoners' behaviour explains that two audits were conducted to evaluate the system's efficiency, and these concluded that there are no biases related to the variables used by the system.

A number of experts stress the importance of impact assessments in relation to fundamental rights. For example, an expert in AI and education states that AI products, especially those making high-risk decisions, require a more robust assessment process than seen in current practices. Three experts in the areas of AI and employment point out that a form of risk assessment related to fundamental rights would be necessary to use AI in recruitment responsibly.

One expert, working for an organisation that helps providers and deployers assess their AI systems, explains that, upon request, they also look into fundamental rights. The organisation for which this expert works can do this based on the Dutch template for impact assessments for human rights in the use of algorithms, but has also developed its own, more simplified, model, which asks questions about 50 fundamental rights. The expert explains that:

> *By asking in advance what the system does exactly and in what context it is being used, certain fundamental rights can be excluded because they are not relevant to that specific system. This makes the process faster and more efficient.*
>
> Expert on assessing AI systems

In addition, this respondent mentions that their company also provides specific tools:

> *… including a light version of the FRIA for systems that are not classified as high risk under the AI Act and do not fall into the prohibited categories [because] a system that is not classified as high risk can still be discriminatory or limit people's autonomy. That's why many clients request additional checks, even though there is no obligation under the AI Act to do so.*
>
> Expert on assessing AI systems

2.3. FUNDAMENTAL RIGHTS RISKS OF HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEMS

The detection of fundamental rights risks is an important first step in addressing such risks, for example to be able to take mitigation measures (see Section **2.4**). The concept of a fundamental rights risk is interpreted broadly as any potential interference with a right ([43]). Providers and deployers can be expected to assess risks to fundamental rights but cannot decide whether there is a violation or not. This decision is ultimately reserved for courts.

The concept of a 'risk-based approach' stems from the corporate responsibility approaches developed by the UN ([44]). It has been adopted in several EU laws that tackle fundamental rights protection in the data and digital fields. This includes the GDPR and the regulation of online platforms through the Digital Services Act ([45]).

This section outlines the potential risks to fundamental rights in the different high-risk areas. This includes both those identified by respondents when discussing current practices and other potential risks to fundamental rights identified in FRA's research. The AI Act does not provide a list of selected fundamental rights that need to be assessed. As FRA has highlighted in a previous report, all fundamental rights may be impacted by the use of AI, depending on the purpose and context of use ([46]). Some fundamental rights are clearly more impacted in certain areas than in others; for example, the right to education (Article 15 of the Charter) could be particularly impacted by AI systems used in the area of education, while the right to asylum and prohibition of *refoulement* (Articles 18 and 19 of the Charter) could be particularly impacted by AI systems used in the area of migration and asylum.

Providers and deployers tend to focus on a narrow set of rights that are potentially impacted, most notably privacy, data protection and sometimes non-discrimination and access to an effective remedy. There is little to no awareness of more specific rights that may come into play in different areas. For example, none of the respondents in education mentions the right to education (Article 14 of the Charter). None of the respondents in employment mentions the freedom to choose an occupation and right to engage in work (Article 15 of the Charter). None of the respondents in law enforcement mentions the presumption of innocence and right of defence (Article 48 of the Charter). None of the respondents in public benefits mention the right to social security and social assistance (Article 34 of the Charter).

This means that, if FRIAs are to be effectively undertaken in practice, more awareness must be created of the range of fundamental rights that may be impacted by AI systems in particular areas and their meaning in the context of AI. This could be done, among other ways, by guiding the attention of those who will be conducting assessments to the rights most likely to be impacted based on a description of the AI system and its context of deployment. Section **3.1** will look more closely at the guidance needed in this regard.

2.3.1. Asylum

In the area of asylum, all providers and deployers mention risks or aspects related to bias and non-discrimination and/or data protection (Articles 8 and 21 of the Charter). Almost all providers and deployers (five out of six) also mention risks in relation to access to an effective remedy (Article 47 of the Charter), highlighting issues such as the importance of explainability and the equality of arms. These particular issues are important, as it can be hard for an applicant's lawyer to get a full understanding of the AI system involved.

We have no feedback from the security authorities; that's our black box. We send our reports there and that's it, that's a one-way ticket … I don't know whether this is somehow regulated by law or something that would make it impossible, but I would be very surprised if the security authorities were to become more talkative.

Deployer of an AI system in the area of asylum

In addition, two deployers using an AI system that supports the search for country-of-origin information mention that applicants for asylum are not informed of the use of the system. One adds that an appeal therefore cannot address the use of AI as such, but instead can only address whether relevant sources have not been considered. One of the respondents compares this with the current situation, wherein parties are also not informed about how sources of information are selected.

Two respondents from providers do not envisage any risks to the right to asylum (Article 18 of the Charter). With prompting, a provider developing a dialect detection system says that they do not think there are such risks, but later indicates that the evidence gathered may have an influence on the decision taken. An expert mentions that, if such a system does not correctly recognise a dialect, an applicant will be presumed to be lying, which will have a negative impact on the prospects of their application. A provider of a system that can help detect security risks states that, as the system does not influence the decision on whether to grant asylum, any impact has been ruled out. A deployer of this system states that individuals 'rarely incriminate themselves'. Based on these statements, it can be deduced that the right to asylum could be impacted, depending on the way these systems' outputs are used.

Experts confirm that the rights to privacy (Article 7 of the Charter), data protection, non-discrimination, asylum and effective remedy could be impacted by AI systems used in migration and asylum procedures. They also mention, without further substantiating why, other relevant rights that may need to be considered, such as the right to dignity (Article 1 of the Charter) and the rights of the child (Article 24 of the Charter).

The recitals of the AI Act mention the right to dignity and the rights of the child ([47]). Both rights could be impacted, depending on the parameters included in an AI system, its invasiveness and the age of the individuals involved. Human dignity entails that people should not be objectivised or dehumanised, for example by questioning someone about their sexual practices or submitting them to 'tests' to establish sexual orientation in the context of asylum proceedings ([48]). When children are concerned, their best interests must be considered based on an individual assessment in the procedure for international protection, which means taking 'due account of the principle of family unity, the minor's well-being and social development – which includes his or her health, family situation and education – and safety and security considerations' ([49]). Any AI system involved should not form a barrier to such considerations being taken into account on an individual basis.

Potential risks to other rights were not mentioned, even though other rights may be impacted by the use cases covered in this report. Examples of such rights include:

- the right to liberty enshrined in Article 6 of the Charter (e.g. if an AI system contributes to decisions to (continue to) detain asylum seekers) ([50]);
- the protection in the event of removal, expulsion or extradition enshrined in Article 19 of the Charter (e.g. if an AI system contributes to decisions to deny applications, relevant considerations for the applicants' protection would need to be taken into account) ([51]);
- equality before the law enshrined in Article 20 of the Charter (e.g. if an AI system contributes to comparable situations being treated differently, such as who is considered to be part of a family unit when determining who qualifies for international protection) ([52]);
- equality between men and women enshrined in Article 23 of the Charter (e.g. if an AI system contributes to gender inequality) ([53]);
- the rights of older people enshrined in Article 25 of the Charter (e.g. if an AI system directly or indirectly discriminates against older people) ([54]);

— the integration of people with disabilities enshrined in Article 26 of the Charter (e.g. if the use of an AI system leads to direct or indirect discrimination against people with disabilities) ([55]);
— the right to good administration enshrined in Article 41 of the Charter and as a general principle of EU law (e.g. if the use of an AI system presents a barrier to accessing documents or giving reasons for decisions) ([56]).

All of the rights mentioned above may therefore be particularly relevant to consider in the context of a FRIA for high-risk AI systems in the area of asylum. **Table 3** provides an overview of possible fundamental rights risks in the area of asylum.

TABLE 3: POSSIBLE FUNDAMENTAL RIGHTS RISKS IN THE AREA OF ASYLUM

Use cases	Risks mentioned by providers and deployers	Other risks mentioned by experts	Other potential risks
AI language assessment for origin determination in asylum procedures	Right to privacy (Article 7), right to data protection (Article 8), right to non-discrimination (Article 21), right to an effective remedy (Article 47)	Right to dignity (Article 1), right to asylum (Article 18), rights of the child (Article 24)	Right to liberty (Article 6), protection in the event of removal, expulsion or extradition (Article 19), equality before the law (Article 20), equality between men and women (Article 23), rights of older people (Article 25), integration of people with disabilities (Article 26), right to good administration (Article 41)
AI system supporting the search for country-of-origin information			
AI system for detecting security-related information in asylum hearings that should be transmitted to security authorities			

Note: All articles cited in this table refer to the Charter.
Source: FRA interviews and research, 2025.

2.3.2. Education

In the area of education, all respondents from providers and deployers acknowledge the potential risks of bias in relation to non-discrimination and almost all (four out of five) mention potential risks to privacy and/or data protection. One provider mentions the problem of 'homogeneous' datasets. Therefore, this provider has curated their data from scratch, enhancing data in relation to under-represented groups, which they call an equity-by-design approach. They mention that this is not the standard for many other companies, as it is cheaper and quicker to use existing datasets.

One deployer acknowledges that the system may rate the reading abilities of children who speak another language at home lower than they are, because they may not be familiar with certain references or words. Because of this, this deployer asked the provider to adjust the texts children are able to read and include the ability to register students learning the language as a second language for better analysis.

Several respondents also mention the risks that an AI system can have for children with disabilities. For example, a system that assesses reading abilities by tracking eye movements poses a challenge for children with dyslexia or children who are partially sighted or blind. One deployer makes a general mention of the issue of gender equality and how boys and girls might perform differently in assessments, without substantiating claims of such potential risks.

The fact of the matter is that people whose first language is Japanese would speak English differently from those whose first language is Spanish and we have to test that accordingly.

Provider of an AI system in the area of education

If you use voice recognition or speech recognition to order a pizza and you get the wrong pizza, it's annoying, but it's not high stakes. If we use speech recognition to score a child's reading fluency attempt, it's really important that we get it right.

Provider of an AI system in the area of education

I would like the final decision to always belong to a teacher, not a digital system.

Rights holder, Sweden

My family comes from [a specific region], where there are many different lovely dialects, but it is very difficult for an AI to determine if such a dialect is right or wrong.

Rights holder, Sweden

Only one provider mentions that their system does not pose any risks to fundamental rights, as it does not replace human judgement. Two providers do not envisage any risks other than those identified in relation to data protection and privacy. With prompting, one deployer acknowledges the relevance of examining the potential impacts of the system on children's rights. They add that some AI systems have built-in protections to safeguard children's rights, such as restrictions against generating harmful material. Only one provider in education mentions children's fundamental rights without prompting.

Other rights are not explicitly mentioned, although several respondents implicitly refer to the potential risks in relation to effective remedy and good administration. For example, a deployer mentions that not all parents/guardians may have received information on the use of the system or, if they did, may not have fully understood it. Another deployer mentions that a letter is provided to parents/guardians stating that the analyses conducted using AI are 'completely objective'. There is no formal complaint system for the outcomes of the system.

None of the respondents mentions the right to education. At the same time, several respondents acknowledge that the system could produce wrong outcomes, leading to incorrect assessments of a child's reading ability. An expert confirms that the biggest risk is that the information that is being provided to the teacher is inaccurate over a long period. This expert adds that it would be very worrying if the data being gathered through the system were used to label children, for example as slow/poor readers or with a specific diagnosis.

This confirms that the right to education could potentially be impacted, particularly if the AI system's assessment contributes to determining the level or type of education that students receive. All the rights mentioned above may therefore be particularly relevant to consider in the context of a FRIA for high-risk AI systems in education. **Table 4** provides an overview of possible fundamental rights risks in the area of education.

TABLE 4: POSSIBLE FUNDAMENTAL RIGHTS RISKS IN THE AREA OF EDUCATION

Use case	Risks mentioned by providers and deployers	Other potential risks
AI system for assessing reading ability in early education	Right to privacy (Article 7), right to data protection (Article 8), non-discrimination (Article 21), equality between men and women (Article 23), rights of the child (Article 24), integration of people with disabilities (Article 26), right to good administration (Article 41), right to an effective remedy (Article 47)	Right to education (Article 14)

Note: All articles cited in this table refer to the Charter.
Source: FRA interviews and research.

Zoom in: selected rights holders' views on the use of AI for assessing reading ability in early education

Of the participants in a focus group and interviews discussing the use of AI for supporting reading development and/or assessing children's reading abilities with six parents of school-age children, only one states that they are aware of such systems. Generally, participants see many benefits of these systems, with some pointing out, for example, that they could help to save time, support teachers in their work, improve individualised support or provide more impartial assessments.

On the other hand, several participants consider that there could be a risk of discrimination. For example, one participant is concerned about the representation and biases in the data used to train AI systems. Another participant expresses concerns about AI being based on large language models and the potential cultural biases in the data used for training. This participant is worried that AI might enforce a standardised way of speaking, which could be different from a local teacher's dialect or a child's background.

Some participants note that a risk may arise if the AI system cannot properly understand different dialects or stuttering. Another participant raises the potential risks of discrimination in systems that analyse reading aloud compared with silent reading.

Participants have mixed feelings regarding risks related to privacy, with one of them stating that it is difficult to imagine data on reading ability being exploited negatively. One participant expresses, for example, concerns about misinformation and the lack of control over the content being taught through AI systems.

Participants emphasise the importance of integrating AI systems into the existing support framework, rather than replacing human interaction and professional judgement. Several participants underline the importance of having enough information available and of involving various stakeholders in the decision-making process to ensure that the use of AI in education is well considered and beneficial.

Source: FRA focus group and interviews.

2.3.3. Employment

In the area of employment, the primary focus of providers and deployers is on the risks of bias in relation to non-discrimination. One deployer explains that the risks of bias can concern many characteristics, from age (which is a protected characteristic in law) to being extroverted or introverted (which is not a protected characteristic in law). Such risks are particularly concerning when AI is used in 'selection scenarios'.

Another respondent explains that, in certain sectors, you can expect biases. For example, they mention that men are more often recommended for jobs in the transport sector and women for those in the healthcare sector, which reflects the current gender imbalance in these sectors. They state that most of the focus is on detecting bias in relation to gender, age and nationality. This may also be linked to the availability of certain types of data, as data on some protected characteristics, such as religious belief or sexual orientation, cannot usually be collected.

No system is unbiased. What is acceptable and how to measure it? I believe that this is important to think about.

Provider of an AI system in the area of employment

I would not like to have a conversation only with a computer.

Rights holder, the Netherlands

AI struggles to capture the feeling of how a person actually is in real life … with the use of AI it is a paper-based assessment of whether or not someone fits a job whereas the recruitment process actually requires an assessment of the personal interaction as well.

Rights holder, the Netherlands

A provider states that scandals related to discriminatory AI systems have increased the focus on the 'fairness' of selection procedures when using AI. The provider does tests to detect bias within their system, with a focus on ethnicity and gender. For example, they use 'nationality' as a proxy to detect biases based on 'ethnicity', because they do not have data on ethnicity 'because this is oftentimes not included in the CV [curriculum vitae]'. They also try to ensure that male and female job titles are treated equally.

Another provider confirms that they are also mainly concerned with biases in relation to gender and ethnicity. However, this respondent also mentions the potential risks for individuals with disabilities, for which their company introduced 'disability-related accommodations in the AI hiring tools' ([57]).

A provider and deployer also refers to the possible risks to privacy and data protection. Another provider refers to privacy as 'a key pillar of the [company's] policy'. They mention that their matching algorithm does not have access to sensitive data and deployers cannot see this information, which is 'a no-brainer'. Both mention the balance between non-discrimination and data protection. The provider and deployer state that 'this creates a paradox: to prove that you're not discriminating on these grounds, you somehow need this data'.

Only one deployer indirectly refers to the right to an effective remedy, mentioning that there is a complaints procedure accessible on the company's website. The same respondent mentions that they had discussions on fundamental rights other than non-discrimination; however, they are not sure which specific rights were discussed.

Potential risks to other rights were not mentioned, most strikingly the freedom to choose an occupation and the right to engage in work ([58]). This right may be impacted, for example, if an AI system used in recruitment directly or indirectly discriminates against nationals from other Member States. Other rights may also be impacted, for example:

- the freedom of thought, conscience and religion enshrined in Article 10 of the Charter (e.g. if the use of an AI system leads to direct or indirect discrimination based on belief or religion) ([59]);
- the right of older people enshrined in Article 25 of the Charter (e.g. if the use of an AI system leads to direct or indirect discrimination against older people) ([60]);
- the rights to information and consultation and the right to collective bargaining and action enshrined in Articles 27 and 28 of the Charter (e.g. if workers are unaware of a system's use or do not receive an explanation of how it works) ([61]);
- the right of access to placement services enshrined in Article 29 of the Charter (e.g. if the use of an AI system is a barrier to accessing free placement services).

All of the rights mentioned above may therefore be relevant to consider in a FRIA for high-risk AI systems used in employment. **Table 5** provides an overview of possible fundamental rights risks in the area of employment.

TABLE 5: POSSIBLE FUNDAMENTAL RIGHTS RISKS IN THE AREA OF EMPLOYMENT

Use case	Risks mentioned by providers and deployers	Other potential risks
AI-based search engine for recommending, ranking or shortlisting job applicants	Right to privacy (Article 7), right to data protection (Article 8), non-discrimination (Article 21), equality between men and women (Article 23), integration of people with disabilities (Article 26), right to an effective remedy (Article 47)	Freedom of thought, conscience and religion (Article 10), freedom to choose an occupation and right to engage in work (Article 15), rights of older people (Article 25), workers' right to information and consultation within the undertaking (Article 27), right to collective bargaining and action (Article 28), right of access to placement services (Article 29)
AI system for supporting hiring decisions by assessing the cognitive, social and emotional competencies of candidates		
AI system for recommending jobs to jobseekers based on their background, preferences and other input		

Note: All articles cited in this table refer to the Charter.
Source: FRA interviews and research.

Rights holders' views on AI in recruitment

A recent Eurobarometer survey on AI and the future of work showed that 57 % of some 26 400 respondents perceived the use of digital technologies, including AI, for selecting job applicants as negative. By contrast, 36 % saw it as positive, while 7 % did not know how they felt. In addition, 50 % of respondents perceived the use of such technologies for gathering additional information on job applicants as negative, while 43 % of respondents perceived it as positive; the rest did not know how they felt.

For the purposes of this FRA report, a focus group and interviews on this topic were held with six jobseekers, who had varying degrees of awareness of such AI systems. The participants generally see the added value in the use of AI in the recruitment process, both for recruiters and, as some point out, for job applicants. Some mention that the use of AI could reduce biases in the process.

Two participants, for example, note that AI might actively support diversity, especially in the way job advertisements are written. One of these respondents suggests that AI could help tailor the language in such postings to be more appealing to a wide audience. The other one suggests that AI could be used to make job descriptions more inclusive and to help prevent unintended bias by replacing gendered pronouns with more neutral options.

On the other hand, the participants also perceive several risks. Most notably, the majority point out the risk that the use of AI could reinforce existing biases and lead to discrimination. There are also some concerns regarding the impact on privacy and data protection. Two respondents, for example, express concerns over data-handling practices and consider that it is essential for applicants to know and decide how their data might be used for AI training.

Overall, participants stress the need for transparency about the role of AI in the recruitment process. They also point out the need for human involvement, considering the irreplaceable value of human judgement in the decision-making process of recruitment.

Sources: European Commission, ***Special Eurobarometer 554 – Artificial intelligence and the future of work: April–May 2024***, Eurobarometer report, 2025, data annex, p. 29 (QB8.2) and p. 28 (QB8.1). For the methodology, see p. 7 of the Eurobarometer report available at the above link. FRA focus group and interviews.

2.3.4. Law enforcement

In the area of law enforcement, the only provider interviewed and one of the two deployers interviewed report being unaware of any particular risks to fundamental rights. The other deployer, namely a deployer of a system assessing a range of risks related to prisoners' behaviour, mentions the risk of biases in relation to non-discrimination, but states that audits have concluded that no such biases exist. Other than that, this respondent notes that they do not think that the tool has any specific impact on fundamental rights.

An expert, who helped develop a system intended to detect false complaints to the police, states that fundamental rights risks are very limited because there is also bias in the identification of false complaints without the use of the AI system. An independent expert states that they do envisage risks for fundamental rights arising from the false complaints system because, for example, the system analyses reports drafted by police officers, who may introduce bias during the drafting process. They and another expert explain that the system may also disproportionately affect individuals with more limited language proficiency, such as tourists, immigrants or individuals from lower socioeconomic backgrounds.

Regarding the system assessing a range of risks related to prisoners' behaviour, an expert mentions that this system analyses factors from the detainees' past, meaning that there may be discrimination based on, for example, socioeconomic background, mental health conditions or intellectual disability. They add that the system appears to systematically assign higher levels of risk to individuals from the Roma population. In terms of nationality, their system, contrary to expectations, assigned lower levels of risk to foreign detainees than national detainees.

None of the respondents mentions privacy or data protection concerns. One expert mentions that the deployer of the system that detects false complaints to the police did not conduct any assessments, as they consider that the system does not process personal data. Another expert disagrees and believes that the system does process personal data.

Respondents only implicitly mention access to an effective remedy, noting that they are not aware of any specific complaint mechanisms and that people have to use existing procedures to complain to the competent authorities. One expert adds that they doubt the police inform individuals about the use of the AI systems. Knowing that an AI system is used and understanding how it may influence decisions is an important prerequisite for access to an effective remedy. In terms of other rights potentially affected, an expert explains that the system assessing a range of risks related to prisoners' behaviour can support decisions that may restrict detainees' freedom of movement, particularly in relation to conditional release, temporary permits and the degree of freedom granted to detainees within detention facilities.

If I didn't even know it [i.e. the AI system] existed, how am I going to know how to complain?

Rights holder, Spain

In the official notification … of why a complaint has been denied or approved, there should be a link where you can really make an appeal if necessary and maybe have a [data protection lawyer], let's say someone who can help you to understand a little bit better …

Rights holder, Spain

Respondents do not mention any potential risks to other rights, most strikingly the presumption of innocence and the right to defence ([62]). This right may be impacted, for example, if an AI system contributes to labelling people as fraudsters or as more likely to commit violence. Other rights may also be impacted by the use cases covered, for example:

- human dignity enshrined in Article 1 of the Charter (e.g. if people are instrumentalised or objectivised by the use of an AI system) ([63]);
- the right to integrity of the person enshrined in Article 3 of the Charter (e.g. if an AI system used to assess risks in relation to prisoners impacts their mental integrity) ([64]);
- freedom of thought, conscience and religion enshrined in Article 10 of the Charter (e.g. if the use of an AI system leads to direct or indirect discrimination based on belief or religion) ([65]);

— freedom of expression enshrined in Article 11 of the Charter (e.g. if an AI system used for surveillance purposes leads to prisoners abstaining from expressing their opinions or peaceful protesting) ([66]);
— equality between men and women enshrined in Article 23 of the Charter (e.g. if the use of an AI system contributes to inequality);
— the rights of older people enshrined in Article 25 of the Charter (e.g. if the use of an AI system leads to direct or indirect discrimination against older people) ([67]);
— the right to good administration enshrined in Article 41 of the Charter (e.g. if the use of an AI system presents a barrier to accessing documents or giving reasons for decisions) ([68]).

All of the rights mentioned above may therefore be particularly relevant to consider in the context of a FRIA for high-risk AI systems in law enforcement. Table 6 provides an overview of possible fundamental rights risks in the area of law enforcement.

TABLE 6: POSSIBLE FUNDAMENTAL RIGHTS RISKS IN THE AREA OF LAW ENFORCEMENT

Use case	Risks mentioned by providers and deployers	Other risks mentioned by experts	Other potential risks
AI system for assessing the likelihood of complaints to the police being false, mostly in relation to insurance fraud in the context of reports of theft and burglaries	Right to non-discrimination (Article 21), right to an effective remedy (Article 47)	Integration of people with disabilities (Article 26), right to liberty and security (Article 6)	Human dignity (Article 1), right to integrity of the person (Article 3), right to privacy (Article 7), right to data protection (Article 8), freedom of thought, conscience and religion (Article 10), freedom of expression (Article 11), equality between men and women (Article 23), rights of older people (Article 25), right to good administration (Article 41), presumption of innocence and right of defence (Article 48)
AI system for assessing a range of risks related to prisoners' behaviour, from the risk of recidivism to the risk of violence towards other prisoners and self-harm			

Note: All articles cited in this table refer to the Charter.
Source: FRA interviews and research.

Zoom in: selected rights holders' views on the use of AI for assessing the veracity of complaints to the police

The use of an AI system for assessing the likelihood of complaints to the police being false was discussed with six rights holders in a Member State where such a system is deployed. These participants have varying levels of knowledge about AI and its use in this context.

Several participants express concerns over the potential bias in such systems, with two pointing out that this could lead to discrimination against certain people or groups. One participant acknowledges that the police may also be biased, but thinks that dealing with a human is still preferable to dealing with a machine. A participant who is overall positive about the use of AI for law enforcement purposes considers that possible biases are introduced by humans, not machines. Another participant is most concerned about the AI tools being used incorrectly.

The discussions highlight that the main risk that the participants envisage is the lack of transparency around the use of AI systems, with some pointing out how their use in decision-making can be problematic. The participants highlight the importance of there being a human in the loop. Some also mention privacy and/or data protection issues, including cybersecurity concerns. Considering the use of AI in law enforcement more broadly, five participants consider it could be beneficial in some cases. For example, one participant considers that it could make the police's work faster.

Source: FRA focus group and interviews.

2.3.5. Public benefits

In the area of public benefits, the majority of respondents from providers and deployers spontaneously mention possible risks of bias in relation to non-discrimination (five out of eight) and privacy and/or data protection (five out of eight). Respondents also occasionally mention the rights of specific groups, sometimes linked to public institutions' broader reporting or other policies in these areas, such as equality between men and women, the rights of the child, the rights of older people and the integration of people with disabilities. For example, respondents mention that a transcription tool may be less accurate when a voice is distorted due to a disability and that digital exclusion may have a greater impact on older people.

Several respondents refer indirectly or implicitly to possible risks to fundamental rights. One respondent from an organisation that is both a provider and a deployer mentions potential barriers to access to an effective remedy for the vulnerable group of people requesting social care services, because they may not have the resources to take steps to remedy negative effects. Another respondent mentions that systems may be too complicated for individuals to fully exercise their rights, which relates to the rights to good administration and effective remedy.

Similarly, another respondent mentions the importance of transparency, as transparency makes it easier to respond to queries and ensure decisions are well understood and justified. They also indirectly refer to the rights to good administration and effective remedy. With prompting about specific rights, one respondent acknowledges that other rights could be impacted, but states that they lack information and knowledge. Just one of the respondents is convinced that their system does not pose any risks to fundamental rights.

There is no mention of potential risks to other rights, most strikingly the right to social security and social assistance. One provider does mention the potential loss of income, which indirectly relates to the right to social security and social assistance.

When AI systems contribute to decisions on whether to grant certain benefits, there is a risk to the right to social security and social assistance, for example if benefits are unjustly denied. Other rights may also be impacted by the use cases covered in this report and other high-risk AI systems being used in public benefits, for example:

— human dignity enshrined in Article 1 of the Charter (e.g. if people are instrumentalised or objectivised by the use of an AI system) ([69]);

— freedom of thought, conscience and religion enshrined in Article 10 of the Charter (e.g. if the use of an AI system leads to direct or indirect discrimination based on belief or religion) ([70]);
— freedom of expression enshrined in Article 11 of the Charter (e.g. if an AI system that is used to assess people's application for benefits causes them to express themselves less freely);
— equality before the law enshrined in Article 20 of the Charter (e.g. if the use of an AI system contributes to comparable situations being treated differently) ([71]).

All of the rights mentioned above may therefore be relevant to consider in the context of a FRIA for high-risk AI systems in public benefits. **Table 7** provides an overview of possible fundamental rights risks in the area of public benefits.

TABLE 7: POSSIBLE FUNDAMENTAL RIGHTS RISKS IN THE AREA OF PUBLIC BENEFITS

Use case	Risks mentioned by providers and deployers	Other potential risks
AI system for supporting the assessment of social insurance applications by determining health-related classifications based on patients' health data	Right to privacy (Article 7), right to data protection (Article 8), non-discrimination (Article 21), equality between men and women (Article 23), rights of the child (Article 24), rights of older people (Article 25), integration of people with disabilities (Article 26), right to good administration (Article 41), right to an effective remedy (Article 47)	Human dignity (Article 1), freedom of expression (Article 11), equality before the law (Article 20), right to social security and social assistance (Article 34)
AI system for transcribing and summarising conversations between applicants for public benefits / services and case workers		

Note: All articles cited in this table refer to the Charter.
Source: FRA interviews and research.

Eurobarometer survey findings on the use of AI in the area of social security benefits

A 2024 Eurobarometer survey on AI and the future of work asked some 26 415 Europeans 'what impact do the most recent digital technologies, including Artificial Intelligence, currently have on social security benefits?'. The survey clarified that 'social security benefits' meant the following: 'accessing information online on people's rights and obligations; claiming healthcare reimbursements, unemployment benefits, maternity benefits or pensions through a phone app or website; having benefits automatically granted, adjusted or removed based on one's life events'.

Overall, 62 % of respondents considered that the impact was positive, while 22 % saw it as negative. Others did not know enough about the most recent technologies (7 %), thought it would depend (3 %) or did not know (6 %).

While this demonstrates an overall positive attitude, it needs to be noted that the question covered a broad spectrum of uses and technologies, from online access to information to the use of new technologies for automatically granting, adjusting or removing social benefits. The use of AI in some of the cases covered presents a higher risk to individuals, as also acknowledged by the high-risk classifications in Annex III to the AI Act. It is therefore important to explore the public's views on these different AI uses in more detail in the future. However, it is important to highlight that, according to these results, more than one in five Europeans think that the use of AI in this area is negative.

Sources: European Commission, ***Special Eurobarometer 554 – Artificial intelligence and the future of work: April–May 2024***, Eurobarometer report, 2025, data annex, p. 5 (QB1.5). For the methodology, see pp. 7 and 12 of the Eurobarometer report available at the above link.

2.4. MITIGATION MEASURES

One core measure that would help to avoid – or would at least reduce – the adverse impact of the development and use of AI on fundamental rights is the implementation of mitigation measures. Mitigation can be understood in line with approaches to corporate responsibility to respect human rights, such as the UN *Guiding Principles on Business and Human Rights*, whereby any potential adverse impacts on human rights need to be prevented, mitigated or remedied ([72]). While prevention and mitigation address adverse impacts before they occur, access to effective remedy must be provided if an adverse impact has occurred.

In this context, mitigation means reducing the (risk of an) adverse impact on or interference with fundamental rights to an acceptable level. That means either not interfering with a right (e.g. not processing any personal data) or reducing the interference to the extent that it is justified (e.g. collecting only personal information that is strictly necessary, such as the person's phone number if they need to be called back, and deleting the number when it is not needed any more). Impact assessments should aim to identify and mitigate risks to prevent violations from materialising ([73]).

Some work on mitigation measures has been already done by different actors ([74]). As mentioned above, violations of fundamental rights (i.e. unjustified interferences) are established by courts. However, providers and deployers of AI systems need to assess risks to fundamental rights by looking at the severity and likelihood of a violation occurring.

Given the complexity of some AI applications, mitigating risks of violations can be difficult, if not sometimes impossible. This is linked to the fact that some circumstances are outside the control of those providing and, even more, deploying AI. For example, if high-quality data are not available and would be too expensive to collect or obtain, it is more difficult and potentially impossible to build an unbiased AI tool. As another example, deployers of AI systems might not have all the insights and knowledge about how an AI system was developed. Providers, on the other hand, might not be able to anticipate all of the modalities of how a system will be deployed in practice. The question of whether the risks can be mitigated to an acceptable level is a central question to ask before any AI system is placed on the market or put into service or deployed in a specific context.

The findings of this report show a fragmented approach to mitigation. What is seen as a mitigation measure to avoid and reduce fundamental rights risks varies among the respondents. In most cases, it remains unclear why certain mitigation measures have been chosen, how effective they are and whether their effectiveness has been studied or tested.

The mitigation measures taken by respondents do not systematically address or fully cover the potential fundamental rights risks identified in Section **2.3**. The focus, in most cases, is on measures that address specific risks to very few fundamental rights, such as risks to data protection and privacy and the risks of bias and discrimination. This is done, for example, by taking data protection measures, such as removing any personal information from data or using synthetic data. Respondents also mention the importance of data quality, such as ensuring representative data is used, carrying out regular audits for bias and using separate datasets for training and testing.

The fieldwork results also show a strong reliance on human oversight as a general mitigation measure. In that context, respondents mention providing clear instructions for deployers and/or training (e.g. with video demonstrations) as measures to make human oversight possible or more effective. The

effectiveness of human oversight as a mitigation measure will depend on how it is integrated into or designed within a system (see Article 14 of the AI Act), whether tendencies to over-rely on the outputs of an AI system (automation bias) are considered and whether it is combined with other mitigation measures. As the European Data Protection Supervisor emphasised in the context of automated decision-making systems, inherent flaws in such systems cannot be fully mitigated solely through human oversight, and the effectiveness of human oversight will depend on the context of use ([75]).

Respondents in the area of employment appear to be the most experienced in and concerned with bias mitigation. They also highlight the challenges related to 'proxies' (see box below on proxies).

The box below provides an overview of the kind of mitigation measures that respondents mention. No structured approach is taken by respondents in any of the areas covered. This clearly calls for more systematic guidance on how to effectively mitigate risks when using AI in high-risk areas.

Bias mitigation is also very complex, because you may well remove bias on one ground but then you will never know what kind of effect this has on other grounds.

Provider of an AI system in the area of employment

If [the AI systems] are being used to make decisions that are high stakes and if the teachers and others are not trained to use them properly, then I think we have … serious issues.

Expert on AI systems in the area of education

Mitigation measures under the AI Act

Providers are required to implement mitigation measures as part of the risk management system for high-risk AI systems under Article 9 of the AI Act. Risk identification, assessment and mitigation thereby need to be in place before the system is placed on the market or put into service, and these need to be continuously reviewed throughout the whole life cycle of the system. The article mandates that providers must identify and analyse reasonably foreseeable risks to health, safety and fundamental rights, as well as implement mitigation measures.

Risks should be minimised through system design, through control measures for residual risks, by respecting transparency requirements and, where appropriate, by providing deployers with training. Moreover, providers are required to test their systems to ensure that their high-risk system performs in compliance with its intended purpose and set requirements. Providers must consider the impact on vulnerable groups, particularly on minors, when managing risks.

Article 27 of the AI Act requires certain deployers of high-risk AI systems to conduct a FRIA. While Article 27 does not explicitly address mitigation measures, it is understood that they form an essential part of the assessment of risks of harm to affected people, as their adoption can reduce the residual risk that the system's use poses to fundamental rights. In addition, deployers must also describe the implementation of human oversight measures and measures to be taken in the case of the materialisation of the risks identified, including the arrangements for internal governance and complaint mechanisms.

Source: Articles 9 and 27 of the AI Act.

Overview of mitigation measures applied as mentioned by interviewees

Risks to privacy and data protection – mitigation measures

- Anonymisation of the training data
- Use of synthetic data for training
- Measures that enable the removal of personal data from datasets *ex post*
- Not collecting any personal data through the AI system
- External audits on data protection compliance
- Collecting only necessary data in line with the principle of data minimisation
- Data localisation
- Consultation of the data protection authority or legal experts

Risks to equality and non-discrimination – mitigation measures

- Ensuring the use of high-quality data that are representative
- Regular updates of datasets
- Use of separate training and testing datasets
- Strict separation of training data and operation data
- Manual selection of training data by staff of the deploying organisation
- Measuring outputs against a human evaluator
- Regular bias testing
- (External) auditing

Cross-cutting mitigation measures

- Engagement of staff in the development and training phase of the system to improve the understanding of its functionalities
- Training of staff on the use of the system and interpretation of results
- Stakeholder consultations
- Consultations of public institutions
- Ensuring adherence with established legal frameworks and industry standards
- Adoption of technical mitigation measures that ensure that users use the system only for the intended purpose
- Not deploying a system until it is fit for purpose
- Publishing information about the system, including the code, to allow for external research
- Providing clear instructions about the system's use, for example through best practices, support, instruction videos and guides
- Clear communication that the final responsibility lies with staff using the system
- Contextualisation of outputs by the users
- Setting and flagging a higher level of false positives to reduce automation bias
- Built-in mechanisms flagging potentially inaccurate results for manual review
- Use of two-eyes principle for confirming results
- Discretion to not use the system when the person responsible considers it would not be reliable
- Documentation checklists to increase the understanding of decisions taken
- Continuous cooperation between the providers and deployers on the performance of the system to meet users' needs

Source: FRA interviews.

Proxies: information indicating protected characteristics

Proxies are seemingly neutral data that nevertheless strongly relate to a protected characteristic under non-discrimination law. For example, postal codes may be a proxy for socioeconomic status or ethnic groups and names can be a proxy for ethnicity or gender.

The processing of proxies could, if an AI system processes such data to the detriment of people or groups with such protected characteristics, lead to discrimination. A few interviewees discuss this issue. A provider of an AI system in education, for example, mentions that they generally do not use proxy data. A deployer of an AI system in employment mentions that, as the address where someone lives can function as a proxy for socioeconomic class, they remove addresses from their data and use distance to the workplace instead.

At the same time, some correlations, such as between the age and the amount of work experience of candidates, are difficult to avoid. An interviewee emphasises the importance of being aware of such proxies in advance. A provider of an AI system in asylum mentions that, due to the strict separation between training data and the operation of the system, the control of proxies is difficult in practice. Their development team cannot verify whether the identification of specific types of information correlates with certain characteristics, as they do not have access to the necessary data.

On the other hand, proxies could be used for bias testing when data on protected characteristics are not available. A provider in employment mentions that they use the proxy of 'nationality' for testing ethnic bias, as data on ethnicity are often not included in curricula vitae. While acknowledging that this is not a very good proxy, they note that it is the best they have.

An expert who focuses on AI in recruitment explains that they try to predict gender based on the available data in the profile of candidates. If they can tell from the remaining data whether someone is female or male, this means that there are still some remaining proxies in the data related to gender. If they cannot, this means that the data can be considered debiased with respect to gender. However, the expert also considers that interventions that aim to address specific biases might unintentionally bring out or even increase other biases, which is especially difficult when considering intersectional groups, such as candidates who are both female and from a migrant background. The complexity of such overlapping factors makes it difficult to ensure fairness for smaller, more granular groups. Moreover, the model might be missing measures to address some specific, perhaps less known, biases.

Sources: FRA interviews; FRA, ***Bias in Algorithms – Artificial intelligence and discrimination***, Publications Office of the European Union, Luxembourg, 2022, p. 24.

It is a good thing that clients start to ask questions about this [risk mitigation measures], but you have certain responsibilities as a company that builds such algorithms.

Provider of an AI system in the area of employment

Endnotes

(43) Malgieri, G. and Santos, C., '**Assessing the (severity of) impacts on fundamental rights**', *Computer Law & Security Review*, Vol. 56, 2025, 106113.
(44) UN, ***Guiding Principles on Business and Human Rights – Implementing the United Nations 'protect, respect and remedy' framework***, Geneva, 2011.
(45) **Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act)** (OJ L 277, 27.10.2022, p. 1.
(46) FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020, p. 96.
(47) **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)** (OJ L, 2024/1689, 12.7.2024, recitals 27, 28 and 48.
(48) Mahlmann, M., '**Human dignity and autonomy in modern constitutional orders**', in: Rosenfeld, M. and Sajó, A. (eds), *The Oxford Handbook of Comparative Constitutional Law*, Oxford University Press, Oxford, 2012, pp. 370–396, at p. 378; judgment of the Court (Grand Chamber) of 2 December 2014, ***A and Others* v *Staatssecretaris van Veiligheid en Justitie***, Joined Cases C-148/13 to C-150/13, ECLI:EU:C:2014:2406, paragraph 65; FRA, ***Current Migration Situation in the EU: Lesbian, gay, bisexual, transgender and intersex asylum seekers***, Publications Office of the European Union, Luxembourg, 2017.
(49) Judgment of the Court (Grand Chamber) of 11 June 2024, ***K and L* v *Staatssecretaris van Justitie en Veiligheid***, C-646/21, ECLI:EU:C:2024:487, paragraph 75. For more information, see paragraphs 80 and 84 of the same ruling; **Charter of Fundamental Rights of the European Union** (OJ C 303, 14.12.2007, p. 1), Article 24(2); Garayova, L., '**Protecting children's rights in the age of digitalisation – Legal implications for the best interests of the child**', *Pakistan Journal of Life and Social Sciences*, Vol. 22, Issue 2, 2024, pp. 7351–7366, at p. 7355.
(50) Beduschi, A., '**International migration management in the age of artificial intelligence**', *Migration Studies*, Vol. 9, Issue 3, 2021, pp. 576–596, at p. 583.
(51) Beduschi, A., '**International migration management in the age of artificial intelligence**', *Migration Studies*, Vol. 9, Issue 3, 2021, pp. 576–596, at pp. 582–583; Kinchin, N., '**Technology, displaced?: The risks and potential of artificial intelligence for fair, effective and efficient refugee status determination**', *Law in Context*, Vol. 37, Issue 3, 2021, pp. 45–73.
(52) Judgment of the Court (Fourth Chamber) of 5 March 2015, ***Copydan Båndkopi* v *Nokia Danmark A/S***, C-463/12, ECLI:EU:C:2015:144, paragraphs 31 and 32; FRA, European Court of Human Rights and Council of Europe, ***Handbook on European Non-discrimination Law***, Publications Office of the European Union, Luxembourg, 2018, p. 146.
(53) FRA, ***Bias in Algorithms – Artificial intelligence and discrimination***, Publications Office of the European Union, Luxembourg, 2022, p. 70; Mendonca de Lima, R., Pisker, B. and Correa, V. S., '**Gender bias in artificial intelligence**', *Journal of Telecommunications and the Digital Economy*, Vol. 11, Issue 2, 2023, pp. 8–30.
(54) Chu, C. H., Donato-Woodger, S., Khan, S. S., Nyrup, R., Leslie, K. et al., '**Age-related bias and artificial intelligence: A scoping review**', *Humanities & Social Sciences Communications*, Vol. 10, 2023, 510.
(55) Land, C. W., '**Disability bias & new frontiers in artificial intelligence**', *The Journal on Technology and Persons with Disabilities*, Vol. 11, 2023, pp. 28–42.
(56) Judgment of the Court (Fourth Chamber) of 8 May 2014, ***H. N.* v *Minister for Justice, Equality and Law Reform and Others***, C-604/12, ECLI:EU:C:2014:302, paragraph 49. This judgment sets out that the right to good administration 'reflects a general principle of EU law' and accordingly, where 'a Member State implements EU law, the requirements pertaining to the right to good administration, including the right of any person to have his or her affairs handles impartially and within a reasonable period of time, are applicable'.
(57) For more on disability-related accommodations in AI hiring tools, see Tilmes, N., '**Disability, fairness, and algorithmic bias in AI recruitment**', *Ethics and Information Technology*, Vol. 24, 2022, 21.
(58) Rezmer, J., '**Work-related human rights and artificial intelligence**', in: Balcerzak, M. and Kapelańska-Pręgowska, J. (eds), *Artificial Intelligence and International Human Rights Law – Developing standards for a changing world*, Edward Elgar Publishing, Cheltenham, United Kingdom, 2024, pp. 214–230.
(59) Temperman, J., '**Artificial intelligence and religious freedom**', in: Quintavalla, A. and Temperman, J. (eds), *Artificial Intelligence and Human Rights*, Oxford University Press, Oxford, 2023, pp. 61–75.
(60) Chu, C. H., Donato-Woodger, S., Khan, S. S., Nyrup, R., Leslie, K. et al., '**Age-related bias and artificial intelligence: A scoping review**', *Humanities & Social Sciences Communications*, Vol. 10, 2023, 510.
(61) Rezmer, J., '**Work-related human rights and artificial intelligence**', in: Balcerzak, M. and Kapelańska-Pręgowska, J. (eds), *Artificial Intelligence and International Human Rights Law – Developing standards for a changing world*, Edward Elgar Publishing, Cheltenham, United Kingdom, 2024, pp. 214–230.
(62) Blount, K., '**Applying the presumption of innocence to policing with AI**', *International Review of Penal Law*, Vol. 92, Issue 1, 2021, pp. 33–48.
(63) Teo, S. A., '**Human dignity and AI: Mapping the contours and utility of human dignity in addressing challenges presented by AI**', *Law, Innovation and Technology*, Vol. 15, Issue 1, 2023, pp. 241–279; Council of Europe, **Recommendation of the Committee of Ministers to member states regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services**, CM/Rec(2024)5, 9 October 2024, paragraph 1.
(64) Council of Europe, **Recommendation of the Committee of Ministers to member states regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services**, CM/Rec(2024)5, 9 October 2024, paragraph 15: 'Under no circumstances should the use of AI and related digital technologies cause intentional physical or mental harm or suffering to a person.'
(65) Temperman, J., '**Artificial intelligence and religious freedom**', in: Quintavalla, A. and Temperman, J. (eds), *Artificial Intelligence and Human Rights*, Oxford University Press, Oxford, 2023, pp. 61–75.
(66) For more on the chilling effects of surveillance, see, for example, Penney, J., '**Understanding chilling effects**', *Minnesota Law Review*, Vol. 106, Issue 3, 2022, pp. 1451–1530; and Murray, D., Fussey, P., Hove, K., Wakabi, W., Kimumwe, P. et al., '**The chilling effects of surveillance and human rights: Insights from qualitative research in Uganda and Zimbabwe**', *Journal of Human Rights Practice*, Vol. 16, Issue 1, 2024, pp. 397–412.
(67) Chu, C. H., Donato-Woodger, S., Khan, S. S., Nyrup, R., Leslie, K. et al., '**Age-related bias and artificial intelligence: A scoping review**', *Humanities & Social Sciences Communications*, Vol. 10, 2023, 510.

(68) Judgment of the Court (Fourth Chamber) of 8 May 2014, ***H. N.* v *Ministry for Justice, Equality and Law Reform and Others***, C-604/12, ECLI:EU:C:2014:302, paragraph 49. This judgment sets out that the right to good administration 'reflects a general principle of EU law' and accordingly, where 'a Member State implements EU law, the requirements pertaining to the right to good administration, including the right of any person to have his or her affairs handles impartially and within a reasonable period of time, are applicable'.

(69) Mahlmann, M., '**Human dignity and autonomy in modern constitutional orders**', in: Rosenfeld, M. and Sajó, M. (eds), *The Oxford Handbook of Comparative Constitutional Law*, Oxford University Press, Oxford, 2012, pp. 370–396, at p. 378.

(70) Temperman, J., '**Artificial intelligence and religious freedom**', in: Quintavalla, A. and Temperman, J. (eds), *Artificial Intelligence and Human Rights*, Oxford University Press, Oxford, 2023, pp. 61–75.

(71) FRA, European Court of Human Rights and Council of Europe, ***Handbook on European Non-discrimination Law***, Publications Office of the European Union, Luxembourg, 2018, p. 146.

(72) UN, ***Guiding Principles on Business and Human Rights – Implementing the United Nations 'protect, respect and remedy' framework***, Geneva, 2011.

(73) Malgieri, G. and Santos, C., '**Assessing the (severity of) impacts on fundamental rights**', *Computer Law & Security Review*, Vol. 56, 2025, 106113.

(74) See, for example, Dutch Ministry of the Interior and Kingdom Relations, ***Impact Assessment – Fundamental rights and algorithms***, The Hague, 2022, Annex 2; and Janssen, H., Seng Ah Lee, M. and Singh, J., '**Practical fundamental rights impact assessments**', *International Journal of Law and Information Technology*, Vol. 30, 2022, pp. 200–232, in particular pp. 220–223.

(75) See also European Data Protection Supervisor, '**Human oversight of automated decision making**', *TechDispatch*, No 2/2025, Publications Office of the European Union, Luxembourg, 2025.

3

HOW TO ASSESS FUNDAMENTAL RIGHTS RISKS OF HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEMS

This chapter outlines what guidance is needed for assessing AI effectively in practice, based on the needs reported by interviewees. The second section in the chapter outlines the main elements needed for the effective assessment of fundamental rights risks.

The time period of the interviews coincided with the early days after the adoption of the AI Act, meaning that its requirements for high-risk AI did not yet apply. As a consequence, in their responses, interviewees across different high-risk areas express a certain level of uncertainty concerning its interpretations, as providers and deployers had not yet developed any procedures for compliance with the act and were also awaiting further guidance.

It's that we're somewhat drowning in all the well-meaning guidelines, regulations, and so on.

Provider and deployer of an AI system in the area of public benefits

More specifically, several respondents highlight issues with the definition of AI (see Section **1.1**) and the classification of AI systems as 'high risk' (see Section **1.4**). Several respondents also express concerns about how regulatory obligations can be met in practice. These insights into current practices and challenges when it comes to fundamental rights protection when using AI, and in view of the existing regulation, inform the actions required and the practical implementation of laws.

Despite the challenges highlighted, there is also a definite undercurrent of optimism among many respondents about what the AI Act could mean for fundamental rights protection and responsible innovation in practice. Several respondents indicate that the regulation will help providers to create better AI though being able to more effectively assess AI risks and by creating a level playing field between providers.

For example, a respondent from a deployer in employment states that, before the GDPR, 'you could notice that there were so many "cowboy" parties who just did whatever they wanted with your personal data' and that the AI Act will have a similar effect to the GDPR of helping to 'separate the wheat from the chaff among AI suppliers'. A provider of an AI system in education, for example, considers that there is an opportunity for the EU to support innovation by providing the necessary tools and guidance to navigate the regulatory framework, which would be especially beneficial for smaller companies with less resources:

> *The AI Act provides an opportunity, for everybody who builds these systems, to pull their socks up and do it properly. I think that's a good thing. I've always been an advocate for the company themselves to take ownership of their [own] ethics basically and I'm, I'm hoping the EU AI Act will make that a regulatory thing. So, it's not that people have to decide to be nice, it means they have to be nice.*
>
> Provider of an AI system in the area of education

3.1. GUIDANCE NEEDS

The findings of this report reiterate that providers and deployers need practical guidance, for example in the form of a template or questionnaire. This includes guidance in relation to the potentially perceived tension between data protection and bias detection, such as guidance on handling data on sensitive characteristics and how to measure bias.

Respondents from providers and deployers in the areas of asylum, education, employment and public benefits indicate that they need guidance on how to assess their systems in relation to fundamental rights. Respondents in law enforcement do not agree on the necessity of a fundamental rights risk assessment, let alone guidance needs in that regard. One expert in AI and law enforcement mentions that providers and deployers would need a protocol on how to evaluate and mitigate risks.

3.1.1. Templates and questionnaires

In terms of the type of guidance, views differ, but a number of providers and deployers agree that a form of template or questionnaire would be useful. One respondent from a deployer in migration states that they would appreciate a checklist or similar tool that guides the deployer through the self-assessment of the fundamental rights risks posed by the system. They state that this would help them to consider risks that they may otherwise overlook.

One respondent from a deployer in education mentions that, currently, they feel they need to 'invent' or put together guidelines themselves based on available information, but they would greatly appreciate ready-to-use support materials that are practical and easy to implement. Another respondent from a deployer in education mentions that having a framework or guidelines to follow would help to ensure that potential issues are identified and addressed appropriately. They also suggest that a valuable support tool that the EU could offer would be an AI service that interviews people conducting risks analyses, asks the right questions, analyses the answers and provides an assessment.

The childcare benefit scandal in the Netherlands involved an AI system, and at no point did people seem to ask whether fundamental rights were being violated. Had they done so, they probably would have concluded 'we're not going to do this'. The fact that there is an AI Act being implemented and discussed now and that large companies will have processes where these issues are addressed, will help in detecting violations of fundamental rights.

Expert on AI systems in the area of employment

Basically, I think it's really good that we have the AI Act. However, as a public authority, we are so overly cautious and restrictive that it might hold us back a bit in the beginning. But in the private sector, I think it's extremely important because those businesses may have stronger incentives to become more efficient, make money, and so on, and are more likely to push boundaries.

Provider and deployer of an AI system in the area of public benefits

I still want to emphasise how important it is that there will be guidelines and support … from the EU centrally because there will be both people who do not want to do this too carefully and organisations that do not have the competence or capacity to do it … When it's so complex, [we need] something that simplifies. It is incredibly valuable and leads to the secure IT [information technology] environment in every organisation in every country throughout the EU. It is incredibly important that here that there are good guidelines.

Deployer of an AI system in the area of public benefits

An expert on AI and employment draws a parallel with the GDPR, suggesting that a similar centralised approach could be beneficial. One respondent from a provider in the area of asylum also mentions needing a template similar to the data protection impact assessment (DPIA), which forces you to think about relevant questions.

Some respondents note that guidance on risk management and assessment in relation to fundamental rights needs to be practical and actionable. One respondent, from an organisation that is both a provider and a deployer in public benefits, states that guidance should not only be aimed at lawyers but also take into account all perspectives within an organisation. An expert on AI and employment mentions that risk assessments should be contextual and that the process needs to be experimental and iterative.

Easy to implement, relatively practical. Not something halfway vague like 'confirmation bias is bad'. That doesn't help me at all, I already know that myself.

Deployer of an AI system in the area of asylum

In terms of practical guidance, one respondent suggests the use of a colour-coding system (e.g. a traffic light system) to indicate the certification and safety of existing large language model systems. This would help the public sector to make informed decisions without having to perform extensive individual analyses.

3.1.2. Training and skills development

Several respondents in education stress the need for training to enable them to carry out fundamental rights risk assessments of AI. A respondent from a provider mentions that risk assessments can be daunting without the right skills. A deployer also stresses the importance of training those involved to provide them with the necessary skills to conduct the assessments. An expert also emphasises the need for training and clearer communication in layperson's terms, as many educators do not understand fundamental rights issues in relation to AI in education.

3.1.3. Bias detection guidance

Several respondents in employment indicate that they need more guidance in relation to the detection of bias. An expert mentions the tension between data protection and detecting bias, as companies need access to sensitive data of individuals, which they cannot be forced to disclose, in order to look for bias within their system. They also highlight the diversity of fairness metrics, which complicates assessment processes.

We have learned that we have to choose a [bias] metric that can easily be explained to non-technical stakeholders, but with 10 different other metrics you can tell 10 different stories.

Deployer of an AI system in the area of employment

A respondent from a provider confirms these difficulties in terms of measuring bias and states that it would be useful to agree on a clear module that explains what is expected. A respondent from a provider also highlights the uncertainty of what 'bias measuring' and 'bias mitigation' mean in practice, among other things, due to the great variety of measurement methods. This respondent adds that they learned that removing bias on one ground may also have an effect on other grounds.

Respondents from both providers and deployers in public benefits also indicate the difficulties of working with sensitive data to detect forms of bias or discrimination. They mention the following reasons for these difficulties: they do not have information on sensitive characteristics, they are not yet aware of how to handle such data or there are difficulties involved in testing. As examples of difficulties involved in testing, respondents mention that a long evaluation period is needed in which end users are aware of their participation. They also mention that sensitive data are not considered suitable for statistical analysis engines and it is therefore difficult to test for bias.

Some respondents also mention other challenges in relation to bias detection. One provider states that it is very difficult to make deployers understand the outputs and possible trade-offs and that good communication between providers and deployers is necessary regarding the data gathered to assess the risk of bias. Finally, one respondent from a deployer in migration mentions that, to fully assess their system for unjustified bias, they would need complete information, including on sensitive attributes.

Several respondents in the areas of migration and public benefits mention a need for further clarity on privacy and data protection aspects. For example, providers in public benefits mention needing guidance on the limits of cloud use, anonymisation and encryption. One respondent from a deployer in migration indicates their need for training on how to handle personal information for users of the AI system.

Key guidance needs

The AI Act includes several provisions that acknowledge the need for the development of more detailed guidance for its implementation, for example concerning the definition of an AI system (Article 96(1)(f)), classification rules for high-risk AI (Article 6(5) and Article 96) and a FRIA template (Article 27(5)).

The responses in this research were gathered at a time when guidance under the AI Act was still under way. The findings of this report show that the following guidance would support the practical implementation of the AI Act and the protection of fundamental rights more broadly.

- There is need for the creation of a systematic evidence base that allows a better understanding to be gained of fundamental rights risks and effective mitigation practices. In particular, further studies and testing of AI systems' compliance with fundamental rights, particularly in high-risk areas, could support providers and deployers in their fundamental rights obligations.
- More guidance is also needed on best practices concerning bias testing.

Source: FRA interviews and analysis of the AI Act, 2025.

3.2. MAIN ELEMENTS FOR EFFECTIVE ASSESSMENT

In the light of FRA's findings in the 2020 report *Getting the Future Right – Artificial intelligence and fundamental rights*, FRA called for mandatory impact assessments of AI systems that cover the full spectrum of fundamental rights ([76]). That report argued that these should be applied before any AI system is used and then regularly repeated.

The impact assessments should take into account the varying nature and scope of AI technologies, including the level of automation and complexity, as well as the potential harm. They should be conducted in a transparent manner. The 2020 report showed that certain fundamental rights are often impacted by the use of AI, notably the rights to privacy, data protection, non-discrimination and access to an effective remedy. These rights are therefore considered relevant for the assessment of all AI systems.

The 2020 report also highlighted the minimum information needed to assess the potential impact of AI on fundamental rights. This includes a description of the purpose and context, the possible harm, the technology used and the accuracy of its outcomes ([77]).

Since the publication of the 2020 report and the more recent negotiations on and adoption of the AI Act, the attention given to the fundamental rights risk assessment of AI has soared ([78]). Many templates for fundamental rights risk assessments have been developed, some of which focus specifically on AI. For example, the Netherlands developed the fundamental rights and algorithms impact assessment (FRAIA) and the Catalan Data Protection Authority developed a FRIA methodology for AI design and development ([79]). As mentioned in Chapter **1**, the Council of Europe Committee on AI adopted a methodology for the risk and impact assessment of AI systems from the point of view of human rights, democracy and the rule of law ([80]). These non-binding tools can already help providers and deployers assess their AI systems.

As explained in Section **3.1**, both providers (Article 9 of the AI Act) and deployers (Article 27 of the AI Act) of (certain) high-risk AI systems listed in Annex III will have to conduct risk assessments from the summer of 2026. For providers, the risk management requirements are laid down in standards. For deployers, the AI Office is tasked with providing a template. In addition, as regards the exceptional use of the otherwise prohibited remote biometric identification systems for law enforcement purposes in publicly accessible spaces, deployers have needed to conduct such an assessment since February 2025, for which the Commission has provided provisional guidance ([81]).

The findings in this report underline that there is still a strong need to translate abstract legal requirements into practical and actionable guidance. This type of guidance is necessary, as the findings in this report show limited awareness and a lack of expertise among providers and deployers about the fundamental rights that may be impacted by their systems ([82]). Providing guidance and examples will help ensure a rights-based, practical and effective approach to the assessment of AI. As mentioned above, a variety of guidance documents (often referred to as methodologies or models, which outline the elements and processes of the assessment) have been made available and are useful to apply ([83]).

The AI Act FRIA for deployers complements the obligations of providers of high-risk AI systems, as providers have limited insight into the specific use of a given AI system. The deployers will therefore build on the information prepared by the providers when completing the AI Act FRIA. Part of the FRIA should therefore be to check whether they received all of the relevant information from providers (Article 13 of the AI Act).

The AI Act FRIA also complements the DPIA, which focuses on risks to the rights and freedoms of data subjects, as many of these systems will also require a DPIA under the EU data protection *acquis*. Deployers should be able to refer to the findings of the DPIA, as relevant, when completing the FRIA, which looks at risks in relation to a broader spectrum of fundamental rights in addition to data protection.

Finally, when conducting an assessment, providers and deployers are encouraged to involve relevant stakeholders in the process. This includes representatives of groups of people likely to be affected by the AI system, independent experts and civil-society organisations ([84]).

The box below sets out the basic building blocks that are relevant when assessing if an AI system is safeguarding fundamental rights. These can be used as a starting point by different actors involved in assessing an AI system.

Building blocks for assessing the fundamental rights impact of AI

Many different approaches have been discussed, proposed, hailed or criticised when it comes to safeguarding fundamental rights when using AI. The following overview provides basic building blocks for any approaches aimed at assessing the fundamental rights risks of AI in a practical, efficient and targeted way. These considerations can also help to highlight opportunities that the use of such technologies can bring for enhanced fundamental rights protection.

To properly address fundamental rights challenges and avoid violations, risks need to be detected, assessed and addressed through mitigation. A final assessment should be undertaken about the extent to which risks are acceptable or not. The following steps are needed as a minimum and should be documented. Their consideration in multidisciplinary teams, with expertise on legal, technical and social aspects, can help to ensure a comprehensive assessment.

Step 1: Describe the system and its purpose

Describe the process and/or the tasks that the system is used for, as well as the technology used, the data that are used to develop the system and further details about its use (including the scale and scope of deployment and the human oversight measures in place).

Step 2: Identify people and groups likely to be affected

Describe who is likely to be affected by the system and how. Consider if any groups in vulnerable situations could be affected, such as asylum seekers, refugees and/or migrants, children, older people, ethnic or racial minorities, those identifying as lesbian, gay, bisexual, transgender, queer, intersex or asexual, people with disabilities and women.

Step 3: Assess, prevent and mitigate the adverse impact on fundamental rights

Identify which fundamental rights, as contained in the Charter ([85]), could be impacted by the use of the system. Start the assessment with the most relevant fundamental rights. Addressing certain highly relevant rights may help address other fundamental rights concerns, as fundamental rights are interrelated and interdependent. Consulting both (representatives of) people affected by the system and experts is crucial to get the assessments right.

As a minimum, always consider the following.

- The rights to privacy and data protection.
  - Considerations.
    - What information about individuals is used and/or produced?
    - Could the use of the system have an impact on individuals' privacy?
    - How do you ensure compliance with data protection law? Consider the legal basis for the data processing and how you ensure that the processing is fair and transparent. In addition, explain how you ensure that the data are collected for specified, explicit and legitimate purposes; are adequate, relevant and limited to what is necessary; are accurate and up to date; and are not kept for longer than necessary. Furthermore, consider what kind of technical and organisational measures you take to ensure the integrity and confidentiality of the data.
  - Mitigation (take preventative and corrective actions).
    - Refrain from actions with any adverse impact on privacy that is not justified (e.g. interfering with intimate situations or communications of individuals). If justified, create safeguards to minimise the adverse impact and test their effectiveness.
    - Refrain from processing any personal data that are not necessary for the purpose of the system. If such data are necessary, justify the use and create safeguards around the processing of personal data, such as limiting their use, storage and access, and test their effectiveness.
  - Assessment.
    - How severe and likely is the impact on these rights in the light of the mitigation measures taken?
- Equality and non-discrimination, including the rights of people with disabilities, older people and children.
  - Considerations.
    - Could the use of the AI system negatively affect certain groups, including by treating them less favourably than others? Consider aspects that make people vulnerable, such as age, ethnic origin, sex, sexual orientation, disability, religion and political beliefs. This may include biased outputs of the system but also, for example, lack of access to a service for certain groups.
  - Mitigation (take preventative and corrective measures).
    - Create a system that does not put certain groups at a disadvantage. If it is not possible to rule out any differences in treatment, make sure they can be justified or mitigated through other means (e.g. providing proper human review or alternative ways of using the service). Test the effectiveness of these measures.
  - Assessment.
    - How severe and likely is the impact on these rights in the light of the mitigation measures taken?
- The right to an effective remedy.
  - Considerations.
    - Do you provide an effective remedy to enable people to complain about the use of the system or its outcomes, including about any harm that occurs when using the system?
    - Could the use of the AI system impact a person's right to go to court and to obtain a remedy?

- Mitigation (take preventative and corrective measures).
  - Make sure to inform people about the use of the system in an understandable and accessible way, also considering applicable obligations under EU data protection law (Articles 12, 13 and 14 of the **GDPR**, Articles 12 and 13 of the **Law Enforcement Directive** and Articles 14, 15 and 16 of the **Data Protection Regulation for EU institutions, bodies, offices and agencies**).
  - Make sure that an effective complaints system is in place, including by informing people where and how they can lodge a complaint.
  - Make sure that appropriate information about how the system works and how the assessments are undertaken is shared with complainants, as well as with oversight bodies when required.
- Assessment.
  - How severe and likely is the impact on this right in the light of the mitigation measures taken?

Assess further fundamental rights relevant for the purpose and context of use (see Section **2.3** for particularly relevant rights in different high-risk areas).

Step 4: Make a final assessment

Considering the above risks of adverse impacts on fundamental rights, is the interference with the rights really needed (i.e. necessary)? In addition, are the risks of rights violations acceptable (i.e. proportionate) in view of the benefits that the use of the system offers in practice (also in view of alternative (not AI-based) ways of achieving the same goal)?

Sources: FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020; FRA analysis of the AI Act, 2025.

Endnotes

(76) FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020, pp. 7–8.
(77) FRA, ***Getting the Future Right – Artificial intelligence and fundamental rights***, Publications Office of the European Union, Luxembourg, 2020, p. 97.
(78) Janssen, H., Seng Ah Lee, M. and Singh, J., '**Practical fundamental rights impact assessments**', *International Journal of Law and Information Technology*, Vol. 30, 2022, pp. 200–232; Mantelero, A., '**The fundamental rights impact assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template**', *Computer Law & Security Review*, Vol. 54, 2024, 106020; ALIGNER, '**Fundamental Rights Impact Assessment**, ALIGNER website, 7 October 2024, accessed 13 October 2025; Malgieri, G. and Santos, C., '**Assessing the (severity of) impacts on fundamental rights**', *Computer Law & Security Review*, Vol. 56, 2025, 106113; Cosentini, A., Pollicino, O., De Gregorio, G. , Ermellino, A., Fontanella, D. et al., '**Assessing the impact of artificial intelligence systems on fundamental rights**', *MediaLaws*, 2025 (see the corresponding annexes on the associated **MediaLaws web page**); Galdon Clavell, G., Azores, M. and Gonzalez, L., ***Impact assessment and auditing framework. Technical Report***, project FINDHR (fairness and intersectional non-discrimination in human recommendation), 2025.
(79) Dutch Ministry of the Interior and Kingdom Relations, ***Impact Assessment and Auditing Framework – Technical report***, The Hague, 2022; Catalan Data Protection Authority, ***FRIA Model: Guide and use cases***, Barcelona, 2025.
(80) Council of Europe Committee on Artificial Intelligence, '**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)**', CAI(2024)16rev2, Strasbourg, 28 November 2024.
(81) **Communication from the Commission – Commission guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)**, C(2025) 5052 final of 29 July 2025, paragraphs 370–377.
(82) See Utrecht University, Rijks ICT Gilde, Straatman, J., Muis, I., Bössenecker, G. et al., ***FRAIA in Action***, 2024. A pilot study using the Dutch fundamental rights and algorithm impact assessment (FRAIA) showed that some of the government organisations that took part in the pilot did not have the right expertise to be able to conduct this type of process.
(83) See, for example, Council of Europe Committee on Artificial Intelligence, '**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)**', CAI(2024)16rev2, Strasbourg, 28 November 2024; Malgieri, G. and Santos, C., '**Assessing the (severity of) impacts on fundamental rights**', *Computer Law & Security Review*, Vol. 56, 2025, 106113; Catalan Data Protection Authority, ***FRIA Model: Guide and use cases***, Barcelona, 2025.
(84) See recital 96 of the AI Act.
(85) This step also includes other applicable law that operationalises these rights, such as international human rights, data protection and anti-discrimination law or case-law.

Conclusions

This report provides an analysis of the current practices in and the challenges faced when assessing fundamental rights implications in the development and use of AI in selected areas. It is based on interviews with providers and deployers of AI systems in the areas of education, employment, migration, law enforcement and public (social) benefits, alongside interviews with experts in these fields. Much of the analysis and many of the views of interviewees reflect the situation with respect to regulatory developments during the fieldwork in the course of 2024. During that year, the EU adopted the AI Act, which is the first comprehensive law that tackles the provision and use of AI.

For the potential of the AI Act to become reality, further guidance is needed on how to assess high-risk AI systems for fundamental rights risks in practice, as called for by a number of interviewees. It is precisely this guidance that will help the potential of the AI Act to be realised in terms of ensuring responsible innovation, fundamental rights protection and a level playing field between providers in practice.

In 2025, discussions took place concerning the simplification of existing EU laws, including in the digital area ([86]). Given that it is still early days in the implementation of the AI Act, simplifying the law may be better addressed by providing sufficiently practical guidance for its implementation. Any simplification considerations need to be based on evidence and informed by the experiences of various stakeholders involved in the AI Act's implementation. Civil-society actors have already expressed concerns about the potential lowering of the fundamental rights protection offered by the AI Act and have called for a focus on its proper implementation ([87]).

This report mainly deals with the duties and responsibilities of those developing and using AI. However, providers and developers being responsible for conducting impact assessments has clear limitations. On the one hand, some providers and deployers may try to avoid applying certain requirements by seeking loopholes and complying with only the absolute minimum requirements. On the other hand, there are limits to their knowledge about fundamental rights risks, due to either an absence of evidence or a lack of legal knowledge. The latter issue can be mitigated to some extent through the provision of practical guidance material.

FRA outputs on AI – including the 2025 report on the **digitalisation of justice** and the forthcoming report on the use of remote biometric identification systems for law enforcement purposes – provide further guidance. However, fundamental rights protection works only in conjunction with proper oversight – a topic that has not been the focus of this report.

In practice, oversight bodies engaged in the protection of fundamental rights and/or overseeing compliance with the AI Act will have to investigate where fundamental rights violations still occur. To ensure proper oversight, sufficient resources, including much-needed technical expertise on AI use, are required. In addition, academia, applied researchers and oversight bodies need to both produce and rely on external studies on the fundamental rights implications of AI.

To ensure that providers and deployers conduct proper impact assessments and to help oversight bodies in their work, further studies and knowledge are required. This includes research into which AI systems can be used with minimal risks and which present real risks to fundamental rights, including when the risk may be too high to use such systems in practice. In addition, studies should look at if and how these risks can be best mitigated in practice.

This report underscores the limits of self-assessment methods, especially in novel areas. Although users are becoming more aware of certain fundamental rights risks associated with AI, both providers and deployers still find it difficult to identify and evaluate those risks. One reason for this is that many of them operate in domains where there are few established precedents. Consequently, they often lack the practical knowledge and experience needed to understand how these systems function in real-world settings.

This limitation also applies to oversight bodies, including those protecting fundamental rights. Those assessing AI, those overseeing its use and those protecting fundamental rights alike need a more comprehensive knowledge base to allow them to make informed decisions about the proper use of AI.

Future efforts should include further cooperation between stakeholders so that they can learn from each other and exchange information. This includes those developing AI, those deploying and using it, those researching and assessing AI systems and those overseeing these systems. The AI Act sets out several ways to bring these stakeholders together, including the AI Advisory Forum, of which FRA is a permanent member and which also includes representatives of academia, civil society and business.

Article 77 of the AI Act empowers public bodies responsible for protecting fundamental rights to access documentation related to AI systems whenever this is necessary for the effective fulfilment of their mandates. This is another important area that needs to be taken seriously by such bodies, including equality bodies and human rights institutions, and requires cooperation and investment.

It is in the joint interest of all stakeholders involved in AI development and use, and the people affected by its use, that the systems work properly. Avoiding the use of AI tools that unnecessarily intrude into the private lives of people or put vulnerable groups at an even further disadvantage benefits all. No one should base their decisions on opaque reasoning that they themselves do not understand.

While the focus on the risks of AI may appear to be a negative approach, it is important, as some interviewees highlight. Conducting proper assessments and ensuring that AI use safeguards fundamental rights are also key to producing better technology that the people affected by it can trust.

Regulation does not hamper innovation, competition or business. On the contrary, it embraces the EU's competitive advantage and high standard of living, which includes respect for fundamental rights.

Endnotes

(86) **Annexes to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Commission work programme 2025**, COM(2025) 45 final of 11 February 2025, Annex I.
(87) European Digital Rights (EDRi), '**Open letter: European Commission must champion the AI Act amidst simplification pressure**', EDRi website, 9 July 2025, accessed 8 October 2025.

Annex: Methodology

This report is based on field research that involved interviews with potential providers and deployers of AI systems and experts on such systems in selected high-risk areas, as defined in the AI Act, namely:

— asylum;
— education;
— employment;
— law enforcement;
— public benefits.

It also includes a limited set of views from rights holders based on either short interviews or focus group discussions with members of the general public.

INTERVIEWS WITH PROVIDERS, DEPLOYERS AND EXPERTS

In total, 38 interviews were conducted with those involved in developing, deploying and assessing AI in selected areas. The interviews covered three respondent groups, which are referred to generically as **interviewees** or **respondents** in the report, namely:

1. **providers** – those from private companies and public institutions that develop AI to place on the market or put it into service;

2. **deployers** – those from private companies and public institutions using AI systems under their authority for professional activities;

3. **experts** – people who formerly worked on the selected AI use cases, academics involved in research on AI, representatives from civil-society organisations working on related areas / use cases and experts from human rights supervisory bodies with expertise on the selected AI use cases.

Sometimes, interviewees were both providers and deployers of an AI system.

Table 8 presents the number of interviews conducted with representatives of each group in the use case areas covered in this report. Overall, the report's analysis is based on 38 interviews.

TABLE 8: NUMBER OF INTERVIEWS PER INTERVIEW GROUP AND USE CASE AREA

Use case area	Number of interviews				
	Provider	Deployer	Provider and deployer (*)	Expert	Total
Asylum	3	3	0	4	10
Education	3	2	0	1	6
Employment	2	2	0	3	7
Law enforcement	1	2	0	3	6
Public benefits	2	1	5	1	9
Total	11	10	5	12	38

(*) Some interviewees were working for entities that were both developing and using AI systems.
Source: FRA interviews and research, 2025.

To obtain information about practices and the degree of awareness of fundamental rights issues when using AI, a separate semi-structured interview took place with each interviewee. The interviews followed the interview guidelines developed for each of the defined interview groups: providers, deployers and experts. The interviewers asked for detailed information about the use of AI, its purpose, its development and its application, and practices for and interviewees' awareness of AI assessments, including with respect to fundamental rights. Due to the limited number of interviews, the views expressed are not representative of all providers and deployers in the areas covered.

The interviews with providers, deployers and experts focused on five Member States: Germany, Ireland, the Netherlands, Spain and Sweden. These were selected to ensure a diverse sample of Member States with respect to the AI policy situation and the level of adoption of AI in the areas covered, and for the methodological feasibility of carrying out the research. The report does not contain a Member-State-level analysis because the areas of high-risk AI covered differed across Member States, which could make such conclusions misleading.

To achieve the highest possible response rates and the best possible insights into the use of AI, the interviews were conducted with a guarantee of interviewee anonymity. This means that the information from interviews is not linked to the people interviewed or their organisations. This step was necessary to get interviewees to agree to the interviews. However, it also limits the information that can be made available based on the analysis, which is necessary to protect anonymity. This means that descriptions of use cases are limited to avoid unique use cases being identified and findings being linked to certain use cases.

FOCUS GROUPS AND INTERVIEWS WITH RIGHTS HOLDERS

This report also includes a limited set of views from rights holders based on either short interviews or focus group discussions with members of the general public. The findings of these interviews and discussions are reported separately in boxes in this report. Overall, a total of 18 people (with an equal gender balance) in the Netherlands, Spain and Sweden either were interviewed or participated in small focus groups to discuss their views on the use of AI in education, employment and law enforcement (six people per area).

LIMITATIONS

Regarding the interviews with providers, deployers and experts, the distinction between providers and deployers was not always entirely clear. In addition, despite efforts to obtain a more gender-balanced sample, it was not possible to obtain a fully equal representation of men and women. This reflects, to some extent, the higher percentage of men working on AI. Overall, 21 (55 %) interviewees were male and 17 (45 %) were female. In addition, 24 (63 %) interviews took place online and 14 (37 %) took place in person.

Given that the total number of interviews held with providers, deployers and experts was 38 – with an additional 18 members of the public taking part in interviews or focus group discussions as 'rights holders' – the findings can claim to provide only a partial picture of what providers, deployers, experts and members of the public think and experience with respect to high-risk AI. Nevertheless, given the paucity of empirical fieldwork-based data collection regarding those developing and using AI, these findings – based on structured in-depth questions – provide valuable evidence. This evidence can inform the application of the law in practice to safeguard fundamental rights.

Background research was carried out to map and learn about different use cases in the specified areas. Use cases were selected based on their relevance and feasibility, giving consideration to how widely they may be applied and their potential risks to fundamental rights. Securing interviewees was challenging, partly because it was not easy to find AI systems that potentially qualify as high risk in the specific areas covered and in the selected Member States. The challenges in securing interviewees may also be related to a lack of trust and openness in the light of current legal uncertainty and the implications of the AI Act.

Getting in touch with the EU

In person
All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you online: (**european-union.europa.eu/contact-eu/meet-us_en**).

On the phone or in writing
Europe Direct is a service that answers your questions about the European Union. You can contact this service:
— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
— at the following standard number: +32 22999696,
— via the following form: **european-union.europa.eu/contact-eu/write-us_en**.

Finding information about the EU

Online
Information about the European Union in all the official languages of the EU is available on the Europa website (**europa.eu/european-union/index_en**).

EU publications
You can view or order EU publications at **op.europa.eu/en/publications**. Multiple copies of free publications can be obtained by contacting Europe Direct or your local information centre (**european-union.europa.eu/contact-eu/meet-us_en**).

EU law and related documents
For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR- Lex (**eur-lex.europa.eu**).

Open data from the EU
The portal (**data.europa.eu**) provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

PROMOTING AND PROTECTING YOUR FUNDAMENTAL RIGHTS ACROSS THE EU

As AI becomes increasingly embedded in daily life, FRA's research warns that AI systems used in high-risk areas, like asylum or public benefits, cannot be used blindly. They should be assessed for their potential impact on fundamental rights. The findings of this report are based on interviews with providers, deployers, and experts working with high-risk AI systems. It finds that awareness of fundamental rights risks and how to mitigate them is limited. This report outlines how the AI Act can be used to effectively protect fundamental rights. It argues that this protection is not a barrier to innovation but instead leads to better-performing technology and greater trust.

FRA – EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS
Schwarzenbergplatz 11 – 1040 Vienna – Austria
Tel +43 158030-0 – Fax +43 158030-699

fra.europa.eu

linkedin.com/company/eu-fundamental-rights-agency
instagram.com/fundamental.rights
facebook.com/fundamentalrights
youtube.com/EUAgencyFRA
x.com/EURightsAgency

Publications Office of the European Union