In this chapter, we aim to analyze the European AI Act in depth. We will explore its scope, delve into its risk assessment mechanisms, and examine the obligations it sets for AI providers and users. By the end of this chapter, you will be equipped to utilize these legal and ethical norms in the real-world scenarios of designing, deploying, and overseeing AI systems. Let's start at the beginning: why did Europe come up with specific legislation for this emerging technology, and what is its connection to the ethics of Artificial Intelligence?

Understanding Europe: The Core of Responsible AI

The European drive to regulate AI was sparked by the ethical concerns prompted by the quick rise of AI in the past years. It is however also part of a broader approach, called the Digital Decade 2030. This approach seeks to transform and digitize the European economy and society. It is part of the so-called European Green Deal, which seeks to transform the EU into a modern, resource-efficient and competitive economy.[1] Both were initiated after the 2008 worldwide crisis. As a result, there has been intense activity in the EU concerning the regulation of digital markets, including the AI sector.

Initial steps towards regulation

The EU first signaled its intentions to regulate AI in a Commission strategy document in April 2018.[2] The three aims of the strategy were to boost the EU's technological and industrial capacity and AI uptake, encouraging the modernization of education and training and ensuring an appropriate ethical and legal framework, based on the Union's values and in line with the Charter of Fundamental Rights of the EU.

In December of that same year, a Coordinated Action Plan followed.[3] Next to various projects and investments, the Action Plan set out how the recently-passed General Data Protection Regulation (GDPR) set rules for use of personal data in AI. An accompanying

By the end of this chapter, you'll be able to ...

- Understand the general structure of the AI Act.
- Analyze the AI Act's risk-based classification and its market implications.
- Summarize relevant requirements for practical AI deployment scenarios.

plan announced the need for ethics guidelines with a global perspective and ensuring an innovation-friendly legal framework.[4] The initial mentions of a need for legislation appears to have been shelved in the meantime.

Work of the High-Level Expert Group on AI

In order to support the development of the AI strategies outlined above, the European Commission established two groups: the High-Level Expert Group on AI (AI HLEG) and the European AI Alliance. The European AI Alliance is an online forum with over 4000 members representing academia, business and industry, civil society, EU citizens and policymakers. The HLEG is a group of experts from industry, academy and society whose work have served as starting points for policymaking initiatives taken by the Commission and its Member States.

In its two-year run, the HLEG has produced these deliverables:

1. *A definition of Artificial Intelligence: Main capabilities and scientific disciplines,* a preparatory document necessary for the other works.
2. *The Ethics Guidelines for Trustworthy AI,* setting an ethical framework for AI based on the notion that trustworthy AI must be lawful, ethical and robust.
3. *The Policy and Investment Recommendations for Trustworthy AI,* a set of recommendations to guide trustworthy AI towards sustainability, growth, competitiveness, and inclusion.
4. *The Assessment List for Trustworthy AI,* a practical tool that translates the Ethics Guidelines into an accessible and dynamic self-assessment checklist.
5. *The report Sectoral Considerations on the Policy and Investment Recommendations,* exploring implementation of the above in three specific areas: Public Sector, Healthcare and Manufacturing & the Internet of Things.

The main deliverable of the HLEG is the *Ethics Guidelines for Trustworthy AI,* an extensive document that systematically analyzes ethical concerns in AI and proposes a concrete framework for addressing them. At the bedrock of this framework is the concept of "Trustworthy AI"; humanity needs to be able to trust the sociotechnical environments in which AI is embedded.[5] Together with the *Assessment List for Trustworthy AI* (ALTAI), the Guidelines form the basis for compliance with the AI Act. This book will use the ALTAI as a central guide throughout the coming chapters.

Towards the AI Act

In October 2019, incoming EC President Ursula Von der Leyen promised to pass AI legislation within her first 100 days in office.[6] The first step was to draw up a White Paper, which outlined three major steps. First, next to new legislation (the AI Act) existing legislation on topics like product liability, ecommerce, motor vehicles would be updated to address applications or impact of AI. Second, all of these efforts would be underpinned by a risk-based approach for new rules. And third, the ethical values inherent in such new rules should be exported to other countries in the world.

The third step in particular underlines the EU's intention to be a leader in AI based on ethical values: "*in a global arena of AI competition and cooperation, the EU attempts to project itself as a Normative Power Europe.*"[7] While this may sound ambitious, it is not unjustified: the so-called "Brussels Effect" of EU regulation having worldwide impact is a proven phenomenon in international politics and economics.[8] As one of the largest and most integrated economies in the world, the EU's stringent regulations often set the standard for global norms. An early example is the 2007 REACH Regulation (Registration, Evaluation, Authorization, and Restriction of Chemicals): chemical companies must identify and manage the risks linked to the substances they manufacture and market in the EU. Its influence has extended beyond EU borders, leading chemical companies worldwide to adopt similar practices to ensure market access in the EU. The 2016 General Data Protection Regulation (GDPR) is also widely cited as exhibiting a similar effect in other countries.

A first draft of the AI Act was released by the European Commission on 21 April 2021. In November 2022, the Council of the European Union (government ministers from each EU country) proposed a set of amendments, followed by the European Parliament's own amendments in May 2023. With their eyes on a January 1, 2024 date of adoption, intense negotiations between the institutes took place in the months that followed.[9] Key concerns were the scope of the so-called prohibited practices, the position of foundation models, (now called general-purpose AI models) the classification as high-risk, and the use of real-time biometrics for law enforcement. On December 8, 2023 a political agreement was reached, which was formalized into the final text in the first months of 2024.

Understanding the AI Act

With over 200 legal clauses and almost 100 introductory recitals providing context and intent, the AI Act rivals the GDPR in complexity and interpretational challenges. Let's dive in and examine the general structure and key terminology of the AI Act.

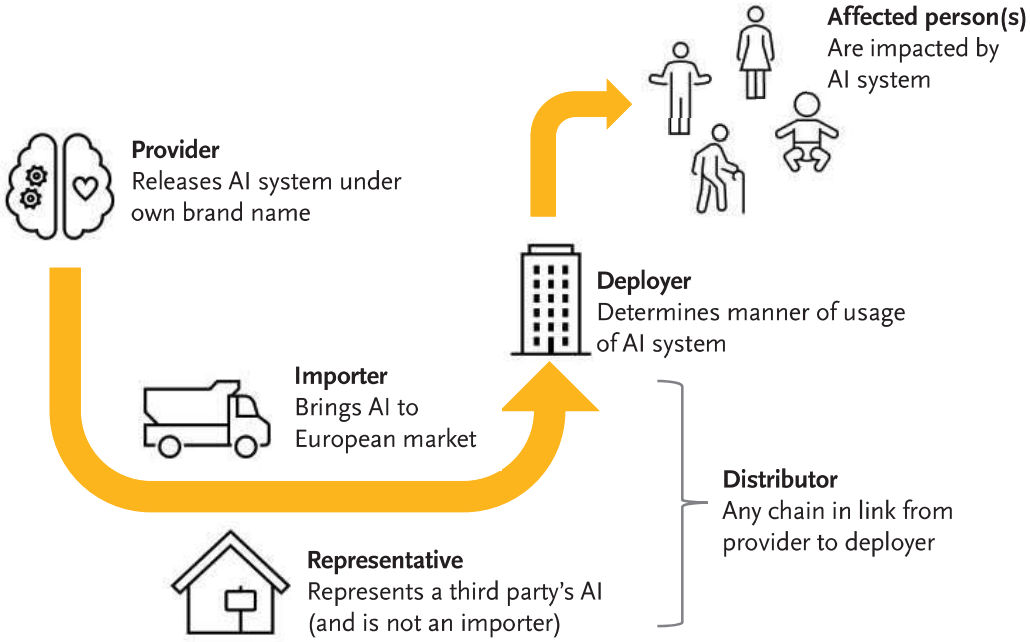
A table of contents

The AI Act counts 83 articles, divided into twelve titles, each of which may contain several chapters.

1. Title I – General Provisions. The first articles define subject matter and scope of the AI Act, including key definitions (discussed in the next subsection) and general tasks for national legislators and supervisory authorities.
2. Title II – Prohibited AI Practices. This title prohibits certain applications of AI because they fundamentally contradict core values of the EU, more on which in the next section on managing risk.
3. Title III – High-risk AI Systems. In this title, divided into four chapters, the AI Act sets out the many compliance requirements and restrictions for so-called "high-risk" AI. These are the subject of the following chapters of this book.
4. Title IV – Transparency obligations. This title defines basic transparency obligations for any type of AI, not just high-risk systems.
5. Title V – General Purpose AI Models. This title defines basic rules for AI models that have multiple applications.
6. Title VI – Measures in support of innovation. The regulatory sandbox is the main focus of this title. We'll meet the sandbox later in this chapter.
7. Title VII – Governance. This title sets up the governance and enforcement structure in the EU, including the AI Office that coordinates activity of the national supervisory authorities throughout the Union.
8. Title VIII – EU Database for High-Risk AI systems. To further stimulate transparency, this title establishes a publicly-accessible database listing every high-risk AI system deployed in the European Union.
9. Title IX – Post-market monitoring and market surveillance. In this title various rules regarding information sharing and market monitoring are established. Of particular note are a duty to report 'serious incidents' regarding the use of AI systems. This title also establishes the investigative powers of the supervisory authorities.
10. Title X – Codes of Conduct. This title establishes a mechanism for self-regulation of AI systems that do not qualify as high risk. These could set requirements for example on environmental sustainability, accessibility, stakeholders participation and diversity of development teams.
11. Title XI – Confidentiality and Penalties. This title sets out the powers to impose penalties, including administrative fines, that supervisory authorities may wield in case of noncompliance. More on this in chapter 10.
12. Title XII – Delegated Acts. This formal title provides the legal basis for the European Commission to unilaterally amend certain parts of the AI Act, notably the list of use cases considered high-risk.
13. Title XIII – Final Provisions. The last title of the AI Act amends other legislation to refer to the AI Act, sets its date of entry into force and introduces an exemption for AI systems already on the market.

Key definitions

The AI Act contains over 50 definitions to ensure consistent application of its provisions. The key terms have been illustrated in the below chart, which visually shows the growth of an AI system from creation to market operation.



The key elements in this chart are:

- As discussed in chapter 1, an **AI system** is a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments.
- A **provider** is a legal entity that develops an AI system and puts this on the market under its own name or trademark. This can be for a fee or at no charge, as with the many open source AI systems available on the internet. A provider may also involve a third party to have the AI system developed.
- The **deployer** of an AI system is the legal entity that actually uses the AI system. A deployer can be a provider itself, but could also license or purchase the AI system from a third-party provider. For instance, an insurance firm may license a SaaS-based AI system to do a first evaluation of claims. The firm would then be the deployer.

- Any entity between provider and deployer is generally called a **distributor.** Formally, it is any entity that that makes an AI system available on the Union market without affecting its properties. If the distributor were to customize or otherwise change the AI system prior to introduction, it would become a provider.
- A specific instance of the distributor is the **importer.** This is the legal entity established in the Union that places on the market or puts into service an AI system that bears the name or trademark of a natural or legal person established outside the Union.
- If an AI system is not actually imported (e.g. as physical product arriving at a harbour or airport), then deployment in the Union requires an **authorised representative** established in the Union who has received a written mandate from a provider of an AI system to assume the responsibilities and liabilities associated with the AI Act for an AI provider. This mechanism is similar to the GDPR's authorised representative for data processing.
- Together, the provider, the deployer, the authorised representative, the importer and the distributor are called the **operator(s)** of an AI system.
- **Affected persons** are those persons or groups who are subject to or otherwise affected by an AI system. This is not necessarily the same as 'users' of the AI system, as one may be affected by an AI system's action without actively using it – or even being aware that this is happening. For example, if a self-driving car operates on the open road, any traffic participant would be an affected person, while only the driver in the car could be called the 'user' of the system.

Material and geographical scope

With the proper terminology established, we can discuss the material and geographical applicability of the AI Act. Material applicability means whether a system qualifies as an AI system, while geographical applicability means whether an action with such an AI system by a particular entity triggers the AI Act's obligations.

For material applicability, first of all the system must meet the definition of an AI system quoted above. But there's more: the AI system must have been 'put into service', which roughly means any supply of the system for distribution or use on the Union market in the course of a commercial activity, whether in return for payment or free of charge. The intent behind this convoluted sentence is to exempt research and non-commercial activities such as open-source development of publicly available AI systems. Unfortunately, many open source initiatives have some commercial aspects, ranging from charging fees for consultancy or support to advertising-driven download pages or soliciting voluntary donations. This will cause significant unclarity down the road.

For geographical applicability, the AI system must have been placed on the market or put into service in a European Union member state (or Iceland, Liechtenstein and

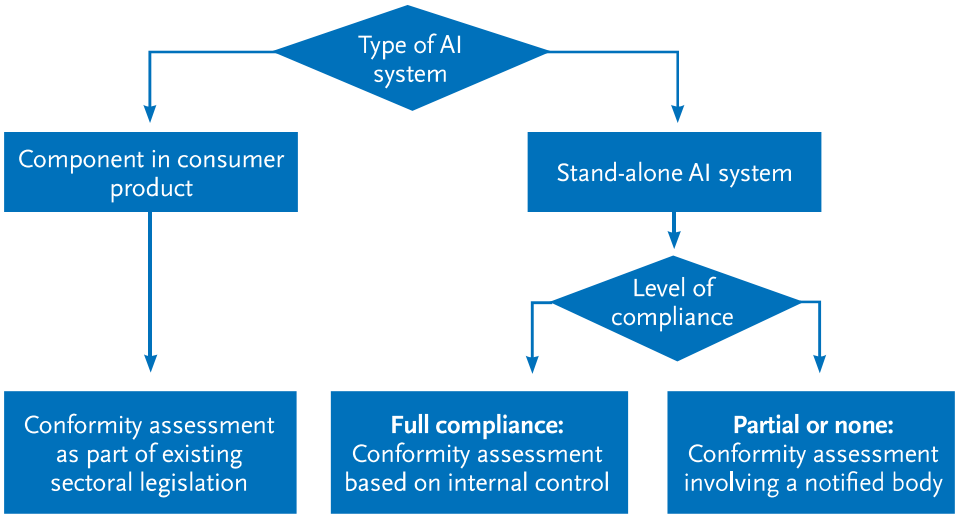
Norway, the European Economic Area members not part of the EU). "Placing on the market" is the EU term for making a product available for sale in a member state. "Putting into service" is the corresponding term for services. Both terms require some form of activity in the member state.

However, there is a broader option: if "the output produced by the system is used in the Union", the AI provider or deployer behind such output is subject to the AI Act. This option thus does not require specific sales or other commercial activities in a member state. If a European firm were to hire, say, a Canadian provider to use an AI system to evaluate the effectiveness of its marketing activities, the output of that system would be used in the Union and thus the Canadian provider would have to comply with the AI Act.

It is irrelevant whether the provider placing the AI system on the market or putting it into service is established within the Union or in a third country. The same goes for providers, importers and distributors. The AI Act thus has a broad reach: a US-based provider that offers access to an AI system through a website is required to comply with the AI Act, despite not having a physical presence in the Union. In practice, the provider would need to appoint an authorised representative who would assume these responsibilities (and face the administrative penalties, including fines, if any violation occurred).

Conformity assessments and certification

The AI Act is written to fit in the existing EU legal framework designed to ensure that products placed within the internal market meet high safety, health, and environmental protection requirements. Central to this framework are the concepts of Notified Bodies and Conformity Assessment Bodies. The latter is an organization authorized to evaluate whether a product, service, process, system, person, or body complies with specified requirements outlined in standards and regulations. A notified body is a specific type of conformity assessment body that has been formally recognized by a national accreditation body to carry out the assessment process for products before they are placed on the European market. In addition, every country has a Notification Authority, which is responsible for ensuring that the Notified Bodies meet the stringent criteria set by the EU to carry out conformity assessments.



High-risk AI systems are required to undergo a conformity assessment procedure prior to their being put on the market or putting into service. The process requires the establishment and examination of a quality management system, specifically tailored to the design, development, and testing of AI products. As illustrated in the flowchart above, the exact process depends on the type of product or service.[10] An AI system embedded in a consumer product already receives a conformity assessment as part of the existing safety legislation for the larger product, and so does not need a separate assessment for its AI component.

For standalone AI systems, the path of a conformity assessment with a notified body is the main option. There is the option of conformity assessment based on internal control, which means the manufacturer themselves assesses and declares that their product meets the relevant EU requirements. This option is only available to a limited subset of the high-risk use cases of Annex III, and then only if the producer is in full compliance with the AIA. All others will have to seek out a notified body, which then conducts an external assessment to ensure the quality management system meets the specified requirements. Any modifications to the system or the AI products it covers must be promptly communicated to the notified body, to allow it to determine if the changes adhere to the established requirements or if a reassessment is necessary.

Floridi et al. have documented example of a conformity assessment procedure for AI.[11] In short, the notified body reviews the AI system's technical documentation, which may include access to training and testing datasets, and potentially the source code, to ensure compliance with the relevant requirements. Failure to provide a comprehensive and accurate application, or to maintain the quality management system to the notified

body's standards, can result in the refusal of the conformity certificate. Such a refusal not only delays the product's time to market but also necessitates a re-evaluation of the system or product, which can be both costly and time-consuming.

In addition to the initial conformity assessment, the European Union employs the instrument of post-market monitoring. This ongoing process requires producers, in this case AI providers, to continuously observe and report on their products' performance once it is introduced to the market. The objective is to ensure that the AI system remains compliant with the necessary standards throughout its lifecycle and that any unforeseen risks or malfunctions are identified and addressed promptly. This proactive surveillance serves to safeguard consumer interests and uphold public safety. However, it also imposes an additional layer of responsibility on businesses, necessitating the allocation of resources to monitor their products post-launch, which can be a significant operational burden, especially for smaller enterprises.

Entry into force and transitional provisions

At the time of writing of this book, the AI Act was scheduled for adoption by the end of April 2024. Like the GDPR, this date will start a two-year transitional period during which AI providers, distributors and deployers will be able to adjust their products, services and processes to the new requirements. However, prohibited practices (see next section) will need to be taken off the market in only six months from the adoption date.

In the meantime, governments can get to work: the AI Act provides them three months to establish notification authorities and notified bodies, required for the certification process, and supervisory authorities that will oversee AI distributors and deployers. They will then have around nine months to produce guidance and procedural regulations for administrative actions and the levels of fines they intend to impose for the various AI Act infractions that may occur.

Managing Risk: AI Practices and Their Classification

The AI Act takes a risk-based approach: rather than specifically prohibiting or regulating certain AI systems or use cases, the Act defines three levels of risk and attaches legal obligations and limitations to each. To understand the implications, let's first take a step back at what we mean with 'risk' in the context of the AI Act.

AI and fundamental rights

The Charter of Fundamental Rights of the European Union holds a distinctive place within the EU's legal framework, specifically in the realm of fundamental rights. These rights are integral to the EU's legal system and take precedence in shaping the Union's laws and policies to uphold human dignity, freedom, democracy, equality, and the rule of law. The Charter's provisions serve as enforceable rights for individuals, and they are pivotal in the EU's legal assessments and decisions.

The terms 'risk' and 'harm' in the context of AI regulation must be understood as a reference to these fundamental rights and values of the Union. As the HLEG put it in the Guidelines: *"Respect for fundamental rights, within a framework of democracy and the rule of law, provides the most promising foundations for identifying abstract ethical principles and values, which can be operationalised in the context of AI."* Deployment of an AI system may introduce a great variety of harms. Here are a few examples of commonly-cited harms, written in terms of fundamental rights with reference to the relevant provisions in the Charter.

- **Privacy and Data Protection (Articles 7 & 8):** AI's capability for mass surveillance and data processing could lead to invasions of privacy, conflicting with the right to the protection of personal data and the respect for private and family life.
- **Non-Discrimination (Article 21):** Biased AI algorithms could result in discriminatory outcomes in services, employment, and justice, which would contravene the principle of non-discrimination.
- **Freedom of Expression (Article 11):** Overzealous AI moderation tools could restrict lawful speech, impinging upon the right to freedom of expression and information.
- **Freedom of Thought, Conscience, and Religion (Article 10):** AI that manipulates information could infringe upon the freedom of thought, conscience, and religion by subtly influencing individuals' beliefs and opinions.
- **Workers' Rights (Article 31):** AI in the workplace could lead to intrusive surveillance and job displacement, undermining the right to fair and just working conditions.
- **Right to a Fair Trial (Article 47):** AI tools used in legal proceedings could lack accountability and transparency, threatening the right to an effective remedy and to a fair trial.
- **Consumer Protection (Article 38):** AI that deceives or manipulates consumers, or that fails to ensure product safety, could violate the right to a high level of consumer protection.
- **Protection of Personal Integrity (Article 3):** AI applications in medicine or biometrics that misuse personal health data or bodily information would conflict with the right to integrity of the person.
- **Environmental Protection and Sustainability (Article 37):** AI systems, through their lifecycle from development to deployment and disposal, can have significant environmental impacts. The energy consumption required for training complex AI

models and the electronic waste generated from rapid obsolescence of AI-enabled devices can contribute to environmental degradation. This poses a risk to the right to a high level of environmental protection and the improvement of the quality of the environment

- **Freedom to Conduct a Business (Article 16):** AI systems could potentially disrupt markets by enabling monopolistic behaviors or by creating unfair competitive advantages through data dominance or algorithmic collusion. This could lead to significant economic harm for smaller companies that cannot compete with AI-enhanced businesses.

There is still significant discussion among scholars how to weigh these harms.[12] Direct physical harm to individuals seems more severe than a long-term negative effect on the environment, for instance. The freedom to conduct a business is limited more easily by courts (e.g. for consumer protection) than the freedom of information. Thus, it is yet unclear which harms should or should not be considered in a risk assessment for an AI system.

Three levels of risk

The AI Act has as its primary aim the mitigation of risks, which it defines as the combination of the probability of an occurrence of harm and the severity of that harm. There is a three-tier approach to managing the aforementioned risks:

1. **Prohibited Practices.** These provide manipulative, exploitative and social control practices that contradict the fundamental rights and Union values, and thus should be banned entirely. Prohibited practices are listed in Article 5 of the AI Act, which means adding or removing such a practice requires a revision of the Act itself.
2. **High-Risk Practices.** While not contradicting fundamental rights or Union values, these practices pose significant risks of adversely impacting these rights. AI systems exhibiting such a level of risk must implement a large number of compliance requirements before being permitted on the market. High-risk use cases are listed in Annex II and III of the AI Act, which the European Commission can amend without revising the Act itself.
3. **Low-Risk Practices.** An AI system that does not pose high risks and does not qualify as a prohibited practice is called a "low-risk" system. These come with basic compliance requirements, mainly focusing on transparency and not taking autonomous decisions.

Prohibited practices

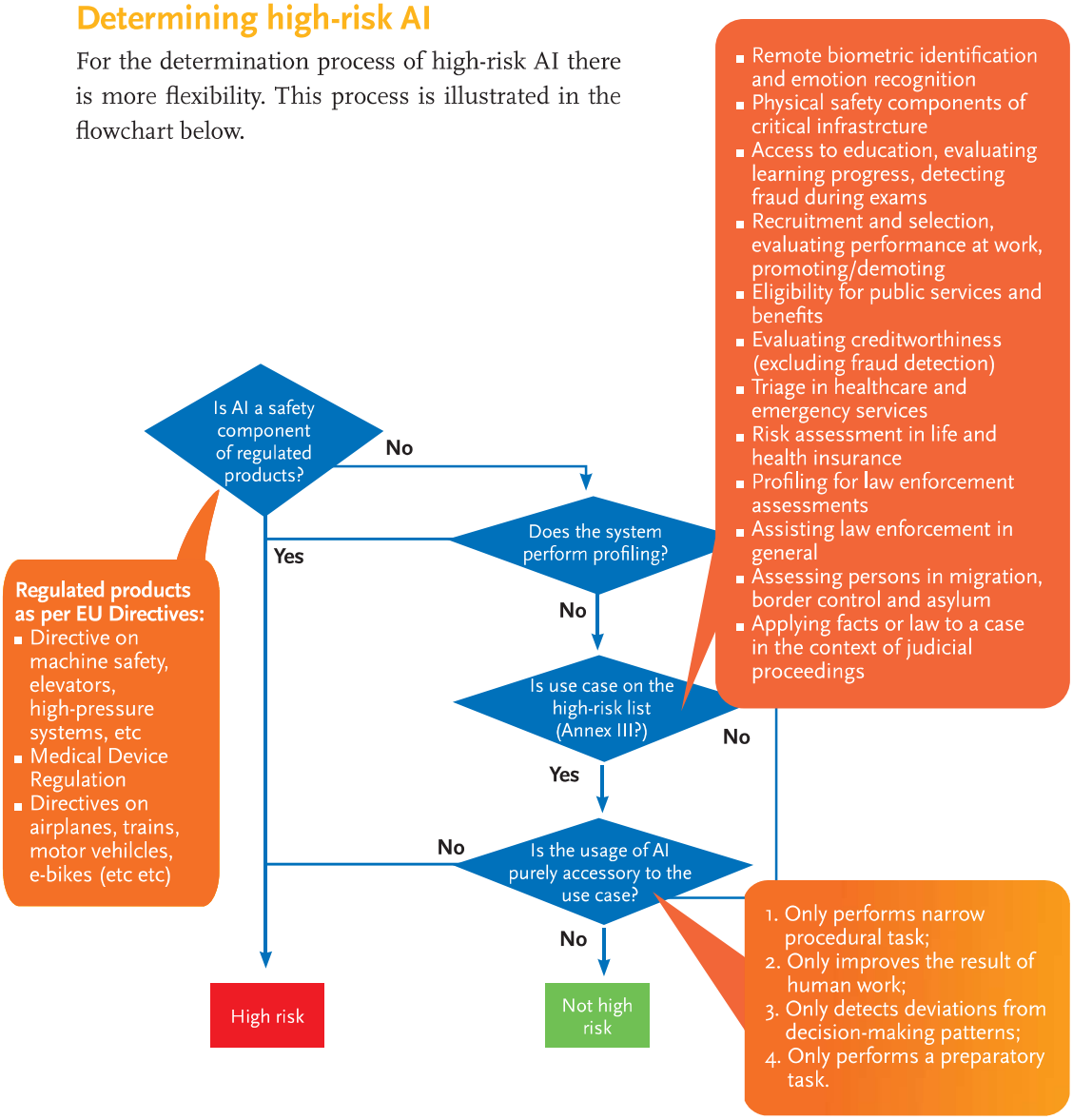
The AI Act prohibits a series of AI-driven practices as being particularly abhorrent: use of AI for "manipulative, exploitative and social control practices" that contradicts Union values of respect for human dignity, freedom, equality, democracy and the rule of law

and Union fundamental rights, including the right to non-discrimination, data protection and privacy and the rights of the child. These practices are enumerated in article 5 of the AI Act and are discussed in more detail in the next chapter.

There is no way to appeal a finding that an AI system is prohibited. The ban is absolute and there are no exceptions. The only real way to avoid the prohibition is to adjust the practice so that it avoids the thresholds of article 5.

Determining high-risk AI

For the determination process of high-risk AI there is more flexibility. This process is illustrated in the flowchart below.



Annex II to the AI Act lists a variety of EU product legislation, such as the safety of production machinery, elevators or toys, the regulation of radio equipment, gas ovens and medical devices. There are also directives on approval and market surveillance of aviation, transport and drones. If an AI system is intended to be used as a safety component of a product, or is itself a product, that falls under any of this legislation, the AI system is by definition high-risk.

More flexibility exists with the use cases of Annex III to the AI Act. Here, many use cases and applications of AI are listed, such as operation of critical infrastructure, recruitment and selection, access to public benefits (e.g. welfare), law enforcement, insurance claim handling and administration of justice. In principle, any use of AI for such an application qualifies as a high-risk case. However, an exception arises when the provider of the AI system can demonstrate that the AI system in fact is "purely accessory" to the use case. An AI system is purely accessory in one of the following four cases:

1. The AI performs only a narrow procedural task, such as automatically adjusting the brightness of streetlights based on environmental conditions.
2. The AI system only improves the result of human activity, e.g. by rewriting a draft into a final text.
3. The AI only detects deviations from decision-making patterns, e.g. by highlighting an outcome that is unusual for the inputs given.
4. The AI system only performs a preparatory task, e.g. classifying or labeling input documents or checking formalities.

To complicate matters, if the AI performs one of these cases but this involves profiling of natural persons (as regulated by the GDPR) the use case is still high-risk.

The onus is on AI providers to justify (with written arguments) why the AI system qualifies as purely accessory. The justification is to be provided to the supervisory authority upon request, who may review and object if the AI is considered misclassified. A fine can in theory be provided for misclassification.

Addressing innovation: regulatory sandboxes

Regulation of new technology poses a dilemma: regulate too early and risk stifling a valuable innovation – but regulate too late and be left with powerless laws against technology giants that have taken over society. The concept of regulatory sandboxes provides a third way: innovators can experiment with new technologies outside existing regulation, while society as a whole remains shielded from any negative impacts.

Origins of sandboxes

Originally developed in the Fintech sector, regulatory sandboxes create a testbed for a selected number of innovative projects, by waiving otherwise applicable rules, guiding compliance, or customizing enforcement.[13] Typically their scope is limited to experiments, as opposed to large-scale production deployments. Despite this limitation, they have long been criticized as being contrary to key principles of law such as legal certainty, proportionality, and equal treatment. The underlying argument is that those that can operate in the sandbox, can achieve a competitive advantage not available to those who operate in the traditional environment, which may seem unfair.

From the beginning, plans for the AI Act had included references to regulatory sandboxes, "for the development, training, testing and validation of innovative AI systems under the direct supervision, guidance and support by the national competent authority". The fear that AI innovation would be stifled was significant, especially given that most innovations in this space already originate from outside the European Union. This is reflected in the stated goals of AI sandboxes: not only should they facilitate training and testing by AI providers, but also provide supervisory authorities with new insights and the ability to draw up guidance for compliance outside of the sandbox.

AI sandboxes in practice

The AI Act does not itself establish sandbox regimes, but rather leaves it to the national supervisory authorities to create appropriate boundaries within European and national law. The modalities and the conditions for the establishment and operation of the AI regulatory sandboxes are to be set down in later legislation, as the AI Act puts it. It is as yet unclear how far this may go: can supervisory authorities waive certain legal requirements or lift restrictions for a sandbox experiment, for instance? Or can authorities go no further than flexible or laissez-faire enforcement of the existing rules?

One point of note is that the AI Act establishes a ground for further processing of personal data for specific regulatory sandboxes, which is a necessity under the GDPR. Other than that, no guidance on sandboxes is available yet.

Related legislation

The AI Act is neither the first, nor the last legislation to address artificial intelligence. It is however the sole European act that specifically regulates this innovative technology. Other laws mainly address outcomes or impact, or indirectly regulate behaviour that may be exhibited by AI systems.

The General Data Protection Regulation

Adopted in 2016, the GDPR is the EU's flagship regulation on personal data, which is a fundamental right in the European legal system. Its focus is on lawful, transparent and fair handling of personal data, which is a much broader topic than just AI systems. However, automated decision-making or profiling of persons has been a key point of attention even before the GDPR: already in the 1990s it was a well-established principle that computers should not exclude people or put them at a disadvantage merely by means of data analysis. The GDPR gave hefty teeth to this principle, with tens of millions in fines available for offenders.

The AI Act and the GDPR obviously overlap where an AI system is trained on personal data or where such a system is used to profile or make decisions. These subjects are handled in chapter 5 (data governance) in more detail. In short, the GDPR takes precedence according to the AI Act. An AI system can also affect persons even when no personal data is processed or when the processing is 'anonymous' in GDPR parlance. In such a case, the AI Act would take precedence.

The AI and Product Liability Directives

The main enforcement mechanism in the AI Act is government intervention by national supervisory authorities. These authorities can issue fines, issue binding instructions or restrictions and even order certain AI practices stopped completely. This very much resembles other legislation, such as the GDPR. However, a key element that is missing in the AI Act is the ability for individual citizens to seek redress from harm caused by AI systems.

This is intentional: the policy choice was made to approach this issue from a product liability perspective. For over a decade, the European Commission has set out to introduce the "New Legislative Framework" to create coherence among the sectorial rules on product safety. In this light, it is understandable to regard risks caused by AI systems as potential defects just like other safety issues with products on the European market. Under the European regime, when safety defects manifest themselves, the producer or importer of the product is liable for any damages unless he can prove the defect could not have been foreseen. (Excluding such liability in terms and conditions is legally not permitted.)

In September 2022 the European Commission proposed the AI Liability Directive, formally called the "Directive on adapting non contractual civil liability rules to artificial intelligence." The gist of the new rules – that have not been formally adopted yet – is that any person harmed by output of an AI system has legal standing to sue its operator. Failure to comply with a relevant duty of care for AI operators makes them automatically

liable for damages, and any noncompliance with AI Act requirements constitutes such a failure. Similar rules are added to the long-standing Product Liability Directive, which creates a similar regime for any form of product safety issue. Of particular importance is the reversal of the burden of proof: once a user of an AI system has made a reasonable argument that the damage is related to actions by the AI-driven system, the provider of the system has to establish that the AI was in fact not at fault.

The two directives are still hotly debated and at the time of writing it is unclear in what form they will be adopted.[14]

Consumer protection and market protection legislation

Many AI systems will be deployed by businesses in interactions with consumers, e.g. to drive sales, improve products or services, engage in customer service or to identify fraud in purchasing transactions. This makes them fully subject to existing consumer protection legislation. Some examples include the Unfair Commercial Practices Directive, the Unfair Contract Terms Directive and the Consumer Rights Directive (CRD). The AI Act declares itself to be complementary to these rules, meaning that even if an AI system is fully brought into compliance with the AI Act it may still violate consumer protection laws.

Most popular AI systems are operated by a handful of very large actors, and made available as internet services. Given their size, they are likely subject to the recently-adopted Digital Services Act and Digital Markets Act. The DSA addresses the legal responsibilities of digital services that act as intermediaries in their role of connecting consumers with goods, services, and content. Such intermediaries face limited liability in their connecting role; the AI Act does not change this. Very large online platforms – service providers with over 45 million European users – face thorough transparency provisions, e.g. on content they remove or users sanctioned for inappropriate behaviour. There are also limitations on profiling, targeting children and other platform actions that may be relevant for AI systems.

The DMA aims to ensure fair and open digital markets. It targets large companies that provide core platform services and have a significant impact on the internal market, serve as an important gateway for business users to reach end-users, and have an entrenched and durable position. The tech giants offering AI for business users, in particular through so-called foundation models, may well qualify as "gatekeepers" under the DMA and face increased scrutiny on their behaviour in the market. One restriction of note is the banning of data sharing between services, meaning an internet platform that offers, say, e-mail and productivity tools cannot use data gathered from user behaviour there to improve an AI system produced as a separate service.

On the topic of (non-personal) data, the upcoming Data Act aims to create a fair and competitive data market, ensuring access to and use of data. For AI producers, this means a more equitable landscape for obtaining the data necessary to train sophisticated AI models. It also imposes obligations on data holders and device manufacturers to provide data access to users, which could increase transparency and potentially lead to more innovation in AI services. Data intermediaries and data sharing for the common good is further regulated in the Data Governance Act, which focuses on the mechanisms of data sharing and governance. It establishes a framework for data intermediaries and mechanisms for data altruism, where individuals and companies can share data for the common good.

European cybersecurity regulations

Recognizing that cybersecurity is a key aspect of the Digital Decade, the European Union has (or is in the process of) adopting a variety of laws regarding cybersecurity requirements. This follows in the footsteps of – again – the GDPR, which already requires personal data processing to be subject to "adequate technical and organization security requirements". An often-heard critique of this legal requirement is that it is too vague and open-ended. The new regulations seek to provide more certainty through standardization and formal assessments.

Already in 2019, the EU adopted its Cybersecurity Act, providing a scheme for voluntary cybersecurity certification, more on which in chapter 4 (robustness and safety). In 2021, the Cyber Resilience Act (CRA) was proposed. The CRA sets cybersecurity requirements for "smart devices", such as internet-enabled music players, robot vacuum cleaners and so on. Many of these devices contain functionality that meets the definition of AI in the AI Act. If the functionality also qualifies as "high-risk", the AI Act prescribes various security requirements, which can be fulfilled by CRA compliance. This in turn requires a specialized audit by an external party.

December 2022 saw the adoption of the "Directive on measures for a high common level of cybersecurity across the Union", commonly known as the NIS2 Directive. This law supersedes the 2016 Directive on Security of Network and Information Systems (NIS) and sets minimum cybersecurity standards for critical or vital infrastructure in both the public and private sectors. This includes telecommunications, transport and banking, but also food, health and the manufacturing of certain economically important products such as medical devices, computer equipment and heavy or advanced machinery. Involvement of AI in such an infrastructure thus requires a careful evaluation of the NIS2 requirements.