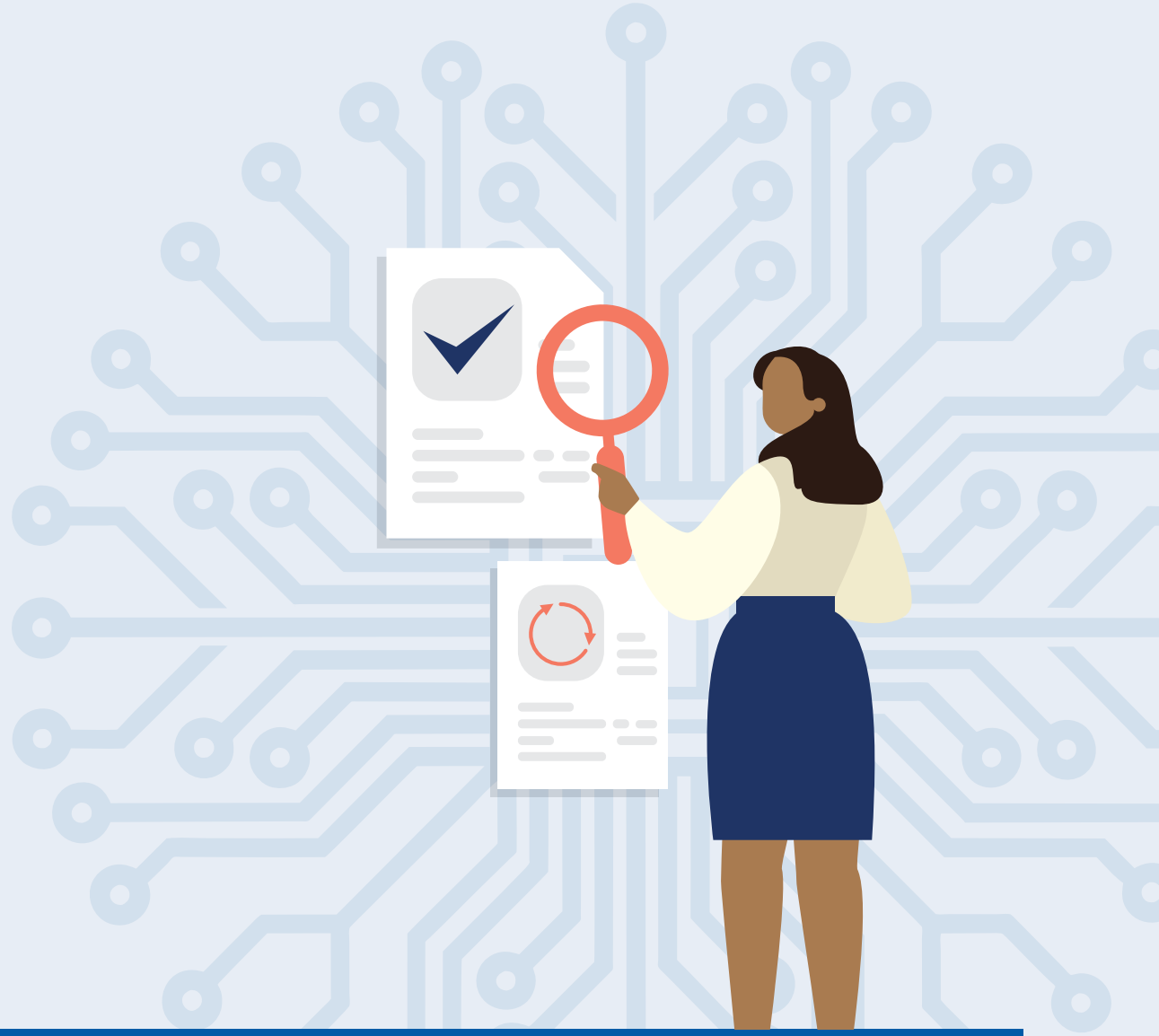
Algorithm Audit



AI AQT documentation

Open source AI and Algorithm Qualification Toolkit for EU and Dutch policy frameworks

September 2026

Table of Contents

1. **Introduction** 3
2. **Questionnaire Identification – AI system** 6
   - 2.1 Inference 7
   - 2.2 Autonomy 7
3. **Questionnaire Identification – Solely automated decision-making** 13
   - 3.1 Decision, legal or similar effects and human involvement 13
4. **Questionnaire Identification – High-impact algorithms** 25
   - 4.1 Direct consequences 27
   - 4.2 Significant effect on the outcome of the process 27
5. **Questionnaire Role and status** 36
   - 5.1 Operator roles 36
   - 5.2 Status of the AI system 37
6. **Questionnaire Risk category** 40
   - 6.1 Annex I – List of Union harmonisation legislation 40
   - 6.2 Biometrics branch 44
   - 6.3 Transparency obligations and prohibited non-consensual fake nude imagery 49
   - 6.4 Annex III high-risk domains 49
   - 6.5 Article 5 horizontal prohibitions 62
   - 6.6 Article 6(3) – Profiling and limited tasks 64
   - 6.7 Article 2 exception 67
7. **Questionnaire Obligations** 69
   - 7.1 Obligations for prohibited AI systems 69
   - 7.2 Obligations for high-risk AI systems 69
   - 7.3 Questionnaire Art. 50 71
   - 7.4 Obligations for generative and interactive AI 80

1. Introduction

The AI and Algorithm Qualification Toolkit (AI AQT) helps with navigating policy frameworks relevant when applying algorithmic systems, such as the GDPR, the AI Act and national policies. The European Union's (EU) General Data Protection Regulation (GDPR) is a cornerstone framework governing how algorithms process data. It regulates how organizations collect, use and share individuals' personal data – whether through analogue means or algorithmic. The EU AI Act has introduced requirements for artificial intelligence (AI) systems to safeguard the safety, health and fundamental rights of EU citizens. In addition, some countries, such as the Netherlands, have implemented broader control measures for algorithmic systems that may have severe impacts on stakeholders. Navigating which legal frameworks apply to a given algorithmic system, and remaining compliant with limited resources, can be challenging. The AI AQT serves as a building block toward compliance across these various policy instruments.

The AI AQT consists of multiple user-friendly, dynamic questionnaires designed to support the harmonized identification and risk classification of algorithmic systems.

Questionnaire **Identification** determines which laws and regulations your algorithmic application falls under:

- **AI system** – As defined in the Article 3 AI Act.
- **Personal data** – Whether the General Data Protection Regulation (GDPR) is applicable.
- **Profiling** – As defined in Article 4(4) GDPR.
- **Solely automated decision-making (sADM)** – Automated decision-making (ADM) practices, including profiling, as restricted or prohibited under Article 22 of the GDPR.
- **Optional:** **High-impact algorithm** – Systems with severe impact on stakeholders monitored by the Dutch government (irrespective of their status under the AI Act). See Algorithm Register Guidelines.

General Purpose AI (GPAI) and requirements for GPAI model providers, as set out in Articles 51-55, fall outside the scope of the AI AQT.

Questionnaire **Role and status** determines what role you have with respect to the AI system and what the usage status is:

- Role:
  - **Provider** – Natural or legal person that develops an AI system or general-purpose AI model and places it on the market under its own name or trademark (Article 3(3) AI Act).
  - **Deployer** – Natural or legal person using an AI system under its own authority in a professional context (Article 3(4) AI Act).
  - **Authorized representative** – A natural or legal person located or established in the Union who has received and accepted a written mandate from a provider of an AI system to carry out on its behalf the obligations and procedures established by the AI Act (Article 3(5) AI Act).
  - **Importer** – Natural or legal person established in the Union who places on the market an AI system bearing the name or trademark of a non-EU entity (Article 3(6) AI Act).
  - **Distributor** – Natural or legal person in the supply chain that makes an AI system available on the Union market without altering its properties (Article 3(7) AI Act).
- Status:
  - **In use** – Must comply by 2030 at the latest with requirements of the AI Act (Article 111 AI Act).
    - General Transparency Obligations (Article 50 AI Act) require compliance by 2 August 2026. AI generated content must be labelled by 2 December 2026

(Article 50 (2)).
    - AI systems that are classified as prohibited AI systems may no longer be used by 2 February 2025. As of 2 December 2026, a new category of prohibited AI systems as introduced by the AI Digital Omnibus becomes applicable. This entails AI systems generating non-consensual intimate images, videos, audio and similar material and child sexual abuse material.
- **In development** – Shorter deadlines depending on the type of AI system (Article 113 AI Act).
    - 2 December 2027 (Annex III – high-risk domains)
    - 2 August 2028 (Annex I – high-risk due to safety component in regulated product)
    - 2 August 2026 (General Transparency Obligations (Article 50 AI Act))

Questionnaire **Risk category** determines the risk classification of the AI system:

- **No requirements** – AI systems with no requirements according to the AI Act (Article 5-6 AI Act).
- **Prohibited AI systems** – AI system meets definition of a prohibited AI system (Article 5 AI Act).
- **High-risk AI systems** – AI system that requires additional control measures (Article 6 and Annex I and III AI Act).
- **Generative and interactive AI** – As described in the Article 50 AI Act (transparency requirements).
- **Exception** – An exception applies for some AI system applications, among others, scientific research, defense, national security etc. (Article 2 AI Act).

The flow of the AI AQT is illustrated in Figure 1. Because the qualification criteria for AI systems, sADM and high-impact algorithms overlap, users are asked a single set of questions in Questionnaire Identification. As high-impact algorithms are particular of interest for Dutch public sector organisations, users can select whether they want this type of algorithms to be identified, or not. Background logic orchestrates the question flow and determines when enough information has been gathered to assess whether a system qualifies as an AI system, sADM and/or high-impact algorithm. Figure 2 presents a Venn diagram illustrating the possible outcomes from Questionnaire Identification and Risk classification. Algorithm Audit has also published supplementary materials on interpreting the AI Act's definition of an AI system[1], as well on the scope of sADM and the requirement of meaningful human intervention.[2] All AI AQT materials on the AI Act are updated according to AI Omnibus amendments.

The AI AQT is designed for organisations seeking a standardized and harmonized approach to classifying algorithmic systems. The tool offers a user interface suitable for both legal expert and users without legal background, translating legal concepts into accessible language.

[1] Implementation of the AI Act – Definition of an AI system, Algorithm Audit (2025).
[2] Meaningful human intervention for risk profiling algorithms – Preventing decision-making based solely on profiling, Algorithm Audit (2025).

Open source code repository

The first version AI and Algorithm Qualification Toolkit (AI AQT) was developed in collaboration with the Municipality of Amsterdam. The tool's source code is available on GitHub and can be (re)used under the EUPL-1.2 license.

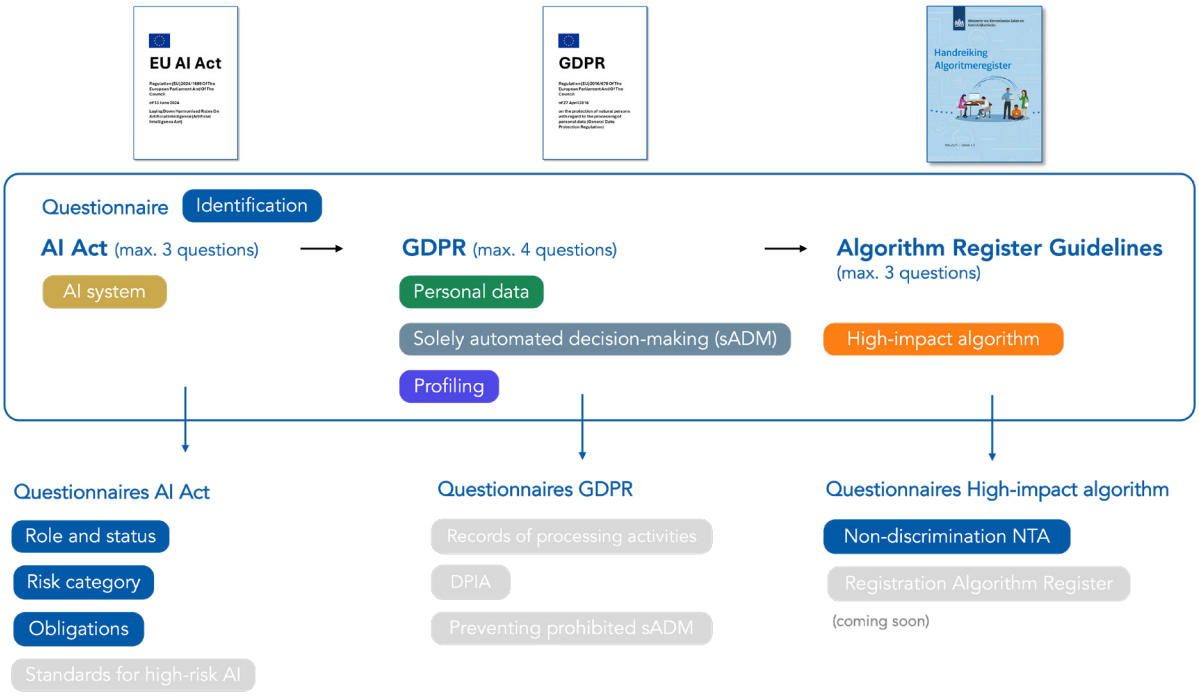


Figure 1 - Overview of the flow of dynamic questionnaires in the AI AQT. Responses to earlier questions are taken into to account in order to minimize the number of questions presented to the user.

The classification of systems into the respective (legal) categories follows a "better safe than sorry" principle, where a conservative classification is preferred over prematurely concluding a system is out of scope. The tool is intended for guiding further compliance actions and does not substitute professional legal assessment or consultation.

This white paper details the considerations and design choices made during development of the AI AQT. It first explains the components of Questionnaire Identification used to determine whether an algorithmic system qualifies as an AI system under the AI Act (section 2). It then addresses GDPR aspects: determining whether personal data is processed, and whether profiling and sADM are at stake (section 3). Next, the concept of high-impact algorithms is introduced and it is discussed how such systems are identified (section 4). Thereafter, the role in relation to AI systems under the AI Act, depending on their usage status, is elaborated on (section 5). Thereafter, it is described how Questionnaire Risk category qualifies AI systems as prohibited or high-risk and whether an exception applies (section 6). The section concludes with obligations that apply for certain combinations of role, status and risk category (section 7).

Box 1

Reserved compliance with the AI Act and Article 22 of the GDPR

This document reflects Algorithm Audit's interpretation of the legal texts of the AI Act and Article 22 of the GDPR, relevant guidelines issued by the European Commission (EC) and European Data Protection Board (EDPB), as well as existing case law. No rights can be derived from this analysis. Ultimately, the competent courts determine the correct interpretation of what constitutes a (high-risk) AI system and sADM.

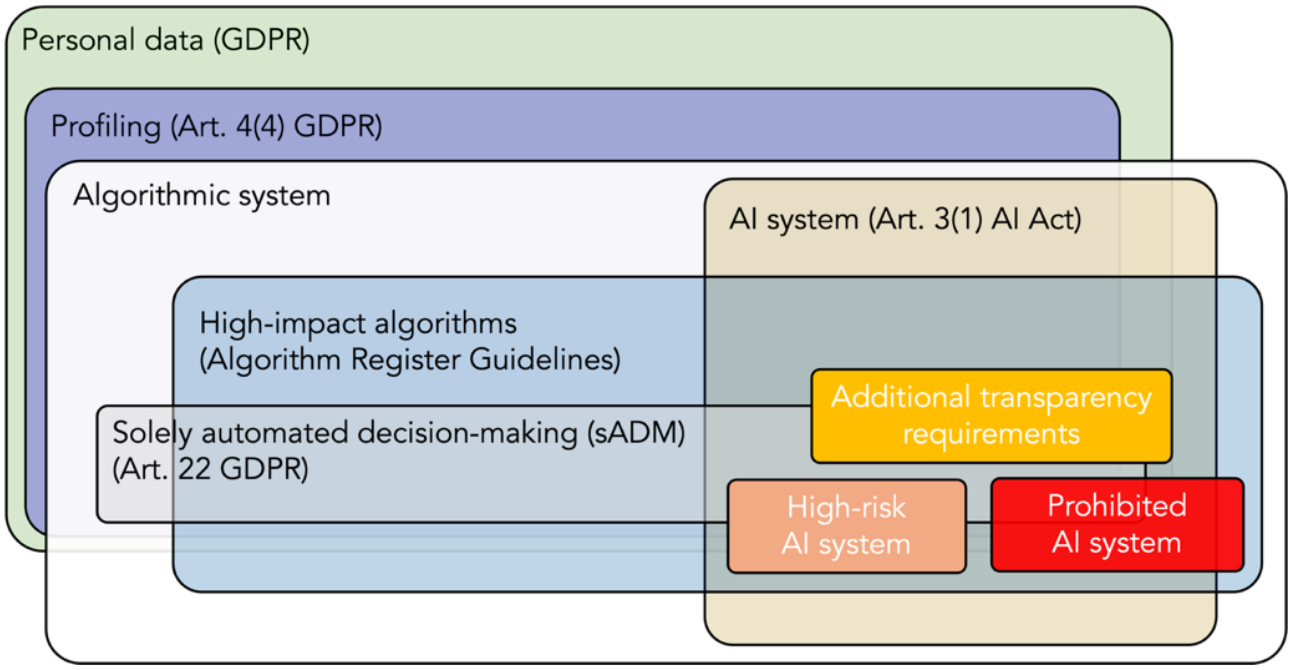


Figure 2 - Overview of the type of algorithmic systems identified by the AI AQT.

2. Questionnaire Identification – AI system

In Questionnaires Identification, Q1-Q3 focus on the definition of an AI system, i.e., the system's design and output. Q4-Q7 determine whether personal data is processed, whether profiling is at stake and whether data is stored and shared. Q8-Q12 determine whether sADM and/or a high-impact algorithm is used.

The definition of an AI system (hereafter: "AI definition") is set out in Article 3(1) of the AI Act. Only systems that meet this definition fall within the scope of the regulation. Article 3(1) defines an AI system as follows:

"… 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."

As elaborated on in the Algorithm Audit's analysis "Implementation of the AI Act – Definition of an AI system"[3], only two concepts in this definition are central to distinguish AI systems from other algorithmic systems:

1. Inference
2. Autonomy.

The other parts of the AI definition are either optional characteristics, such as 'adaptiveness', or concepts that do not contribute to differentiating AI systems from other IT systems, such as 'machine-based'.

This section begins with an explanation of what is meant with inference (2.1) and autonomy (2.2). Guidelines provided by the EC (hereafter:

[3] Implementation of the AI Act – Definition of an AI system, Algorithm Audit (2025).

'guidelines') on the interpretation of the AI definition are incorporated in this analysis.[4] Next this section discusses how these concepts are incorporated into the AI AQT Questionnaire Identification as Questions 1-3 (Q1-Q3).

Algorithm Audit has published examples and explainers clarifying what type of outputs, inference and rule-bases systems fall within, and which systems fall outside the scope of the definition.[5]

2.1 Inference

The capability to infer is the most important element of the definition that distinguishes AI systems from other algorithmic systems. Recital 12 of the AI Act states that: *"A key characteristic of AI systems is their capability to infer"*. The guidelines state that inference is an *"indispensable condition"*. Inference means both the capability to derive models or algorithms from data and to derive outputs from input. Recital 12 further clarifies that (at least) two techniques enable inference: machine learning and logic- and knowledge-based approaches.[6]

There are many techniques which are not traditionally considered to be machine learning, such as statistical approaches where model parameters are 'fitted' on data. These techniques can still be considered inference from data, they are thus included in the AI definition. This is captured Q2.

Logic- and knowledge-based approaches to AI are different. Logic- and knowledge-based approaches are usually carefully manually designed. Although with these approaches there may not be a model derived from data, they are still considered as techniques that enable inference for their reasoning capacity. The question whether the application is a logic- and knowledge-based system is included in Q3.

The AI definition mentions "predictions, content, recommendation, or decisions" as forms of output of an AI system. Because deriving these types of outputs is conditional to conclude that the system has the capability of inference, the output type in included in Q1. Because the target audience of AI AQT – developers, product owners, line managers and other executive users – are typically familiar with the output of their application, Questionnaire Identification starts with asking about the output an algorithm produces.

More information on the concept of inference can be found in section 3 of Algorithm Audit's analysis "Implementation of the AI Act – Definition of an AI system".[7]

2.2 Autonomy

The one other concept in the definition required to distinguish AI systems from other algorithmic systems – aside from inference – is autonomy. Recital 12 of the AI Act indicates this to mean: *"AI systems are designed to be operate with varying levels of autonomy, meaning that they have some degree of independence of actions from human involvement and of capabilities to operate without human intervention"*. 'A certain degree' is however a weak requirement: a system does not have to be completely autonomous to meet this requirement. The guidelines mention that a system which generates an output by itself from manually provided in puts is already considered 'some degree of independence'.

This implies that every system that derives outputs itself is autonomous to a certain extent. In other words, if the inference requirement is met, the autonomy requirement is also met. We conclude that, in conjunction with inference, the 'autonomy' requirement imparts no meaningful criteria to

4 'Guidelines on the definition of an artificial intelligence system established by AI Act', European Commission (2025).
5 Examples and explainers AI Act, Algorithm Audit (2025).
6 Recital 12 of the AI Act.
7 Supra note 3.

distinguish AI systems from other algorithmic systems.

Identification Q1 – What type of output does the application derive?

The output generated by an algorithmic system gives an indication whether it qualifies as an AI system. Users of AI AQT are therefore first asked to indicate what type of output is generated by the system. See Figure 3.

As discussed in 2.1 Inference, deriving outputs is considered to be integral to the AI definition. Types of outputs should be predictions, content, recommendation, or decisions.

Prediction is a broad concept with differing interpretation across domains. In data science, a prediction does not have to be about the future. It can also relate to a data point that has not been observed before. In fact, every score, ranking, recommendation, label, classification, decision and generated content (image, text, speech etc.) is a prediction. Therefore, the answer options include explanatory terms for various types of predictions that should be recognizable to users (score, ranking, label, object-, face- or voice recognition). This choice favours accessible language. If one of these options are selected, users are forwarded to Question 2 (Q2). Note that multiple options can be selected.

Because dashboards are a common type of data-driven application, confusion about whether they qualify as an AI system can arise. On their own, dashboards only provide data visualization. No inference is at stake and they do not transcend *"the elementary processing of data by enabling learning, reasoning or modelling"*[8] as an AI system would. In and of itself, a dashboard cannot constitute an AI system, even if its coupled to one. If a user indicates the only output of a system is a dashboard, it is concluded it is not an AI system. Users are explicitly prompted to consider if other types of outputs are displayed in this dashboard. If they select one of the specified outputs along with "Dashboard", they are brought to Q2.

[8] Recital 12 of the AI Act and section 3 of Implementation of the AI Act – Definition of an AI system, Algorithm Audit (2025).

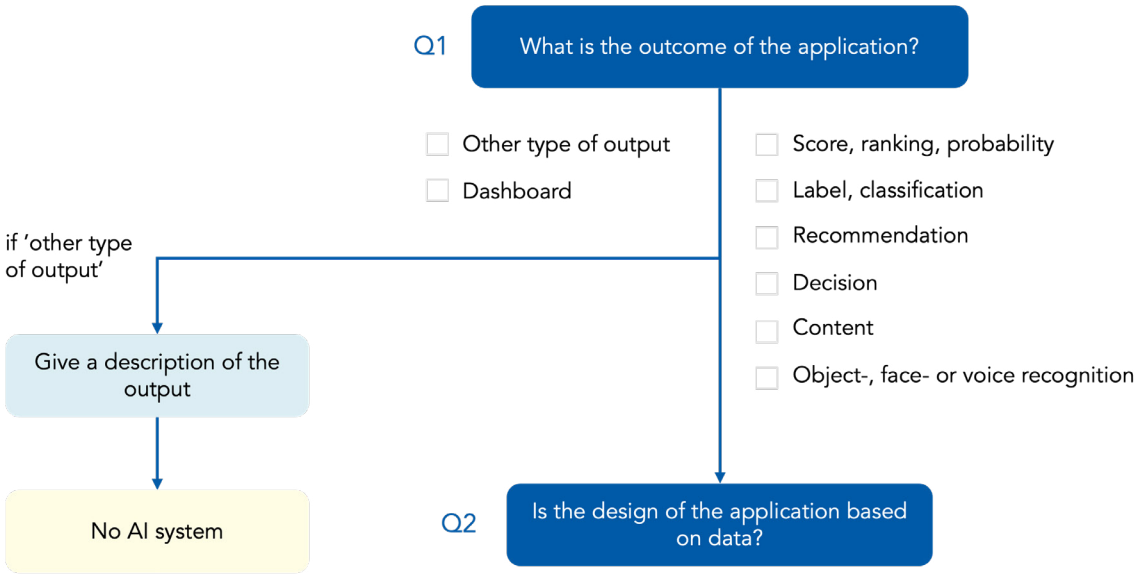


Figure 3 - Identification Q1 asks what type of output the algorithmic application generates.

The same logic applies to the option "Other type of output". If it is the sole provided answer, the tool concludes the application is not an AI system. If coupled with another output (other than dashboard), the user proceeds to Q2. In either case, the user is asked to provide a description of the output, which can be manually assessed by experts.

Identification Q2 – Is the design of the application based on data?

While Identification Q1 focuses on the system's output, Q2 examines how the output is generated (see Figure 4). Under the AI Act, an AI system is defined by its capacity to perform inference. As discussed in 2.1 Inference, determining whether the design of a system relies on data is a key factor in assessing whether inference takes place. This aspect is incorporated into Q2.

If the application contains components derived from data, then it is an AI system. This is the case, for example, when a model or algorithm is learned or fitted using statistics, optimization, simulation or machine learning or a similar technique. In this case, the conclusion – that the algorithm qualifies as an AI system – is presented. The user is then asked whether they want to continue with the rest of Questionnaire 1, which deals with 3. Questionnaire: Solely automated decision-making and 4. Questionnaire: High-impact algorithms, starting with Q8.

Recital 12 of the AI Act indicates that not all applications whose components are derived from data qualify as an AI system. Where design choices are made manually and data analysis is used only to inform those choices, the resulting algorithmic system does not constitute an AI system.[9] When this answer is selected, users are asked to provide an explanation. This information supports a case-by-case assessment, based on expert knowledge, of whether the algorithm does qualify as an AI system. After the clarification question, users are asked whether they want to continue with the rest of Questionnaire 1, which deals with 3. Questionnaire: Solely automated decision-making and 4. Questionnaire: High-impact algorithms, starting with Q8.

[9] This exception is discussed in detail in the paper "Implementation of the AI Act – Definition of an AI system". Supra note 3.

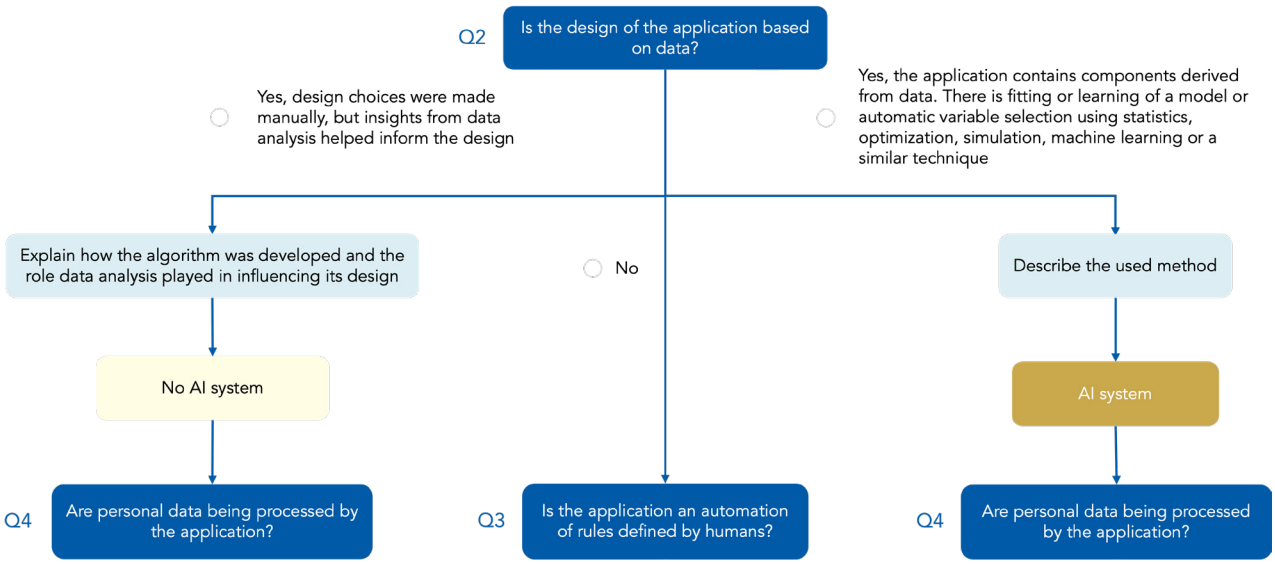


Figure 4 - Identification Q2 is about whether the design of the application is based on data.

Even if the design of the application is not based on data, the application can still be an AI system. To check if this is the case, users are redirected to Q3.

To assist users, the following remark is provided:

Data includes all forms of electronic information. Text, images, and audio are also data.

Applications can be designed manually. However, even when they are manually designed, the design is sometimes based on data analysis. For example, threshold values for (eligibility or exclusion) rules may be calculated from data, or criteria may be selected on the basis of calculated correlations.

In other cases, components (e.g., models and algorithms) are derived more automatically from data. This may involve, for example, fitting a statistical model to data or using machine learning to train a model or rule-based algorithm from data. Simulation and optimization techniques may also be used to derive a model from data.

Large language models such as ChatGPT are also derived (trained) from large amounts of textual data.

Identification Q3 – Is the application an automation of human-defined rules?

Q3 serves to capture a specific case in which no model is derived from data, but the system still qualifies as an AI system, following recital 12 of the AI Act and the guidelines provided by the EC. See Figure 5.

As discussed in 2.1, there exists a case where the design of a system is not based on data but there is still inference involved - that of case of logic and knowledge-based systems. These are often understood or referred to as rule-based systems that involve a certain level of manual programming. For this reason Q3 deals with how the rules a system follows are designed.

Logic- and knowledge-based approaches are explicitly stated in the AI Act as techniques that enable inference and thus should be considered AI. In practice these techniques are usually applied in

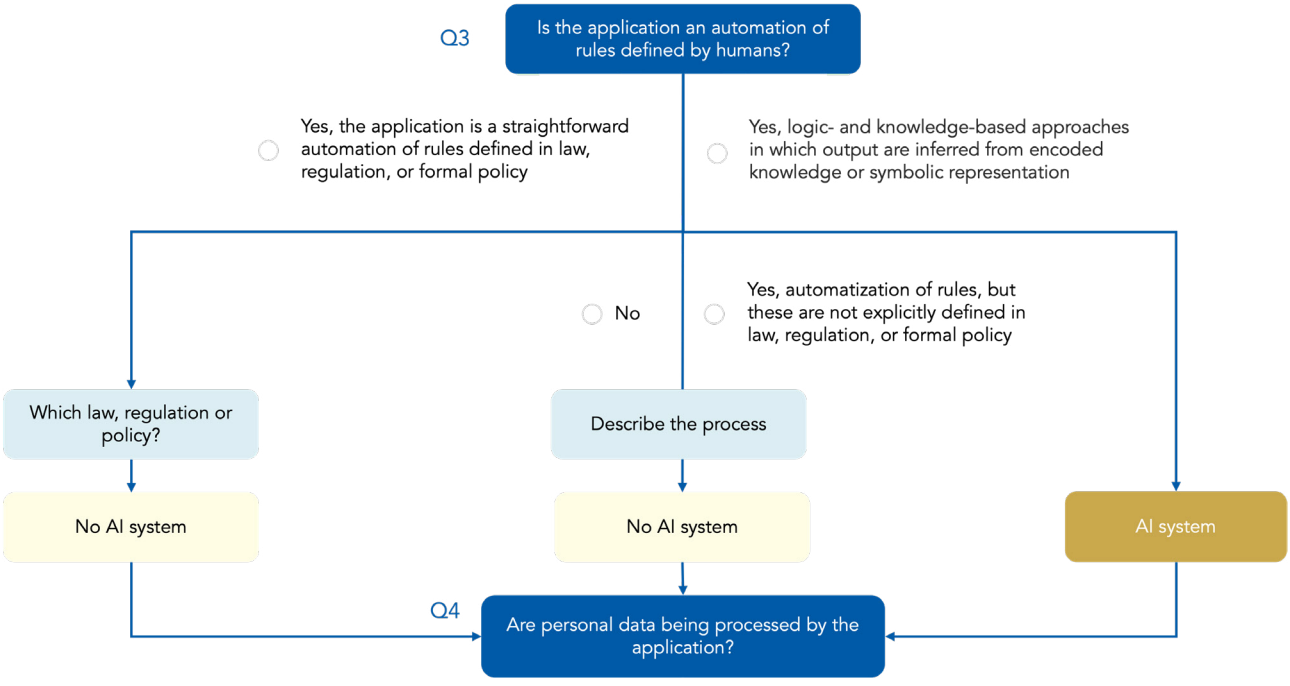


Figure 5 - Identification Q3 examines the extent to which there is human involvement in the creation of rules used in an algorithm.

conjunction with models derived from data (as already captured by Q2).[10] Purely logic- and knowledge-based approaches are rare. Developers using this type of technology are aware this represents a highly specific type. In this case the user is informed the application qualifies as an AI system. Then, users are asked whether they want to continue with the rest of Questionnaire Identification, which deals with high-impact algorithms and sADM, starting with Q8.

In case an algorithm is a straightforward automation of rules defined in law, regulation or formal policy, it does not qualify as an AI system, since there is no inference at hand. In this case, the user is first asked which official law, regulation or policy is automated, before the conclusion that no AI is involved is shared. The input provided by the user can be checked by legal experts.

In case the application is an automation of rules that are not explicitly defined in law, regulation or formal policy, the user is asked to describe how the rules were determined, before the conclusion that no AI is involved is raised again. The input provided by the user can be checked by legal experts. There is likely no inference at play. However, the distinction between rules that are explicitly stated in law, regulation or policy, and those that reflect a human interpretation or implementation of the above, is important for qualifying high-impact algorithms and sADM, as discussed in sections 3 and 4.

To assist users, the following remark is provided along Q3:

An example of rules laid down in legislation or regulations is a rule-based algorithm that, when an application for social benefits is submitted, automatically indicates whether the income and other requirements have (not) been met. In that case, the rules in the algorithm are a one-to-one implementation of norms specified in, for instance, the Dutch Participation Law.

When a standard is defined in open terms in legislation or regulations and is further specified in the application, the application does not constitute one-to-one automation of legislation or regulations.

Examples of rules defined by humans include:

> A rule-based algorithm in which a work instruction has been implemented into an algorithm;
> A risk profile in which the rules have been manually defined on the basis of employees' experience;
> Open legal standards that are further specified in rules.

Logic- and knowledge-based approaches are also referred to as 'symbolic AI systems'. This category of AI systems includes knowledge representation, inductive (logical) programming, knowledge bases, inference and deduction engines, and (symbolic) reasoning. This technology is used, for example, in expert systems.

[10] See section 3.2 of Implementation of the AI Act – Definition of an AI system, Algorithm Audit (2025).

Flowchart identification AI system

Algorithm Audit

Flowchart – AI system (Art. 3 AI Act)

This schematic representation shows the logic required to determine whether the application qualifies as an AI system according to Article 3 of the AI Act. The flowchart of the complete Identification questionnaire with all paths and outcomes can be found on the Algorithm Audit website. The complete questions can be found in the AI AQT tool itself.

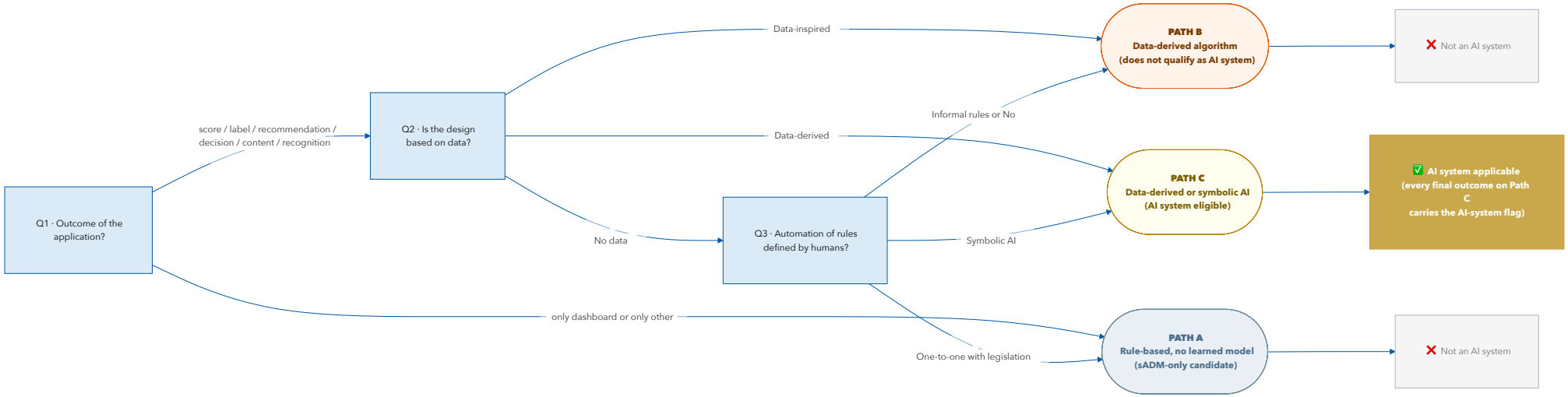


Figure 6 - Flowchart determining whether an AI system is applicable or not.

3. Questionnaire Identification – Solely automated decision-making

Article 22(1) of the General Data Protection Regulation (GDPR) states that:

"The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."

Various resources help to clarify the scope of solely automated decision-making (sADM), including the relevant exemptions set out in article 22(2) GDPR. Competent authorities and legal experts have issued guidance how requirements can be met, such as *"Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679"*[11] issued by the European Data Protection Board (EDPB). In addition, several other sources were consulted in designing the questionnaires used in the AI AQT, including the *"Advice on Article 22 GDPR and automated selection tools"*[12] published by the Dutch Data Protection Authority (DPA), "Advice on automated selection techniques" by legal experts at Pels Rijcken[13] and relevant legal-scientific literature[14],[15]. Algorithm Audit has also published supplementary materials providing a step-by-step guide how decision based solely on automated processing can be prevented.[16]

This section first examines the core concepts that serve to identify algorithmic systems that are in scope of article 22 of the GDPR. In particular, the notion of a 'decision', 'legal effects or other significant impacts for individual' and the understanding of 'based solely on automated processing', including the latter's relationship to human intervention, are discussed (3.1). As will be shown the notions of 'decision with legal or similarly significant effect' and 'automation' from article 22 strongly overlap in meaning with those of 'direct consequences' and 'significant effect on the outcome of the process' from the Algorithm Register Guidelines for Dutch public sector organisations. As such, the same questions can be used to qualify high-impact algorithms and sADM (i.e., Q9 and Q12 in the Identification questionnaire). Hence, section 3.1. explicitly links to concepts explained in section 4. Questionnaire: High-impact algorithms.

NOTE: Not all instances of sADM qualify as a high-impact algorithm. For example, one-to-one automation can qualify as sADM but is strictly outside the scope of high-impact algorithms (see Q3 and 4.2 Significant effect on the outcome of the process).

3.1 Decision, legal or similar effects and human involvement

The scope of the prohibition in article 22 GDPR depends on at least three aspects: I. A decision is made, II. The decision has legal effects or otherwise significantly affects the individual concerned, and III. It is based solely on automated processing. Each concept is discussed in turn.

[11] Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, European Data Protection Board (2018).

[12] Advice on Article 22 GDPR and automated selection tools, Dutch Data Protection Authority (2024).

[13] Advice on automated selection techniques, Pels Rijcken (2024).

[14] Legal protection against risk profiling based on the GDPR, the ECHR, and the Charter of Fundamental Rights, F. Çapkurt, Dutch journal for legal professionals (2025).

[15] The Right to an Explanation in Practice: Insights from Case Law for the GDPR and the AI Act, L. Metikos en J. Ausloos, Law, Innovation and Technology (2025).

[16] eaningful human intervention for risk profiling algorithms – Preventing decision-making based solely on profiling, Algorithm Audit (2025).

I. A decision is made

The notion of a decision must be interpreted broadly: not only formal decisions, as defined in Dutch public administration Law (Awb Article1:3) may affect citizens and organisations. For any type of decision-making, the individual impact must be considered (see II). The Schufa case further broadened the scope by concluding that even the computation of a score itself can constitute a decision.[17]

II. The decision has legal effects or otherwise significantly affects the individual concerned

A decision, which the output of the algorithm informs, has a 'legal effect' or otherwise 'significantly affects' individuals if one of the following types of decisions is made:[18]

- **i.** A formal decision, such as imposing a tax assessment, granting or denying a benefit or allowance, making a decision following an appeal, or granting or denying a permit or subsidy;
- **ii.** A decision with financial consequences, such as the ability to obtain a payment plan or qualify for credit;
- **iii.** Entering into an agreement, such as an employment contract or a purchase agreement;
- **iv.** Selection for an inspection, if the inspection is intrusive for the individual, such as a home visit;
- **v.** A decision affecting someone's access to education, such as admission to a university or school assignments;
- **vi.** Decisions affecting someone's employment opportunities, such as processing job applications or assigning projects to freelancers;
- **vii.** Otherwise significantly impacting the individual.

The types of decisions listed above correspond to the answer options included in Q5. It is worth noting that inspections, in particular, can have a significant effect before even concluding or without ever resulting in further actions towards an individual beyond the inspection itself. For example, an inspection can serve as grounds to delay payments or temporarily suspend access to services. A particularly intrusive inspection, like a home investigation, can have material consequences for the subject, even if not financial. Alternatively, a history of inspection can legitimatize increased caution in future cases and provoke a sequence of checks. As such, decisions to subject an individual to an inspection can be highly impactful in and of themselves and warrant detailed attention. For this reason a supplementary question, Q5.1, is raised to ascertain the secondary effects of an inspection-related decision when Q5 is answered with "Inspection, investigation, or requests for additional information".

Secondary effects might also occur if the outcomes of a risk profiling algorithm shared internally or externally, or are stored long-term significant consequences for stakeholders can arise, as discussed under the explanation for Q8-Q10 hereunder.

More information on 'legal' or 'similarly significant' effects can be found op p.21-22 of the EDPB guidelines and p.6-7 of the advice published by the Dutch DPA.

III. Solely based on automated processing

A system falls within the scope of article 22 of the GDPR when decisions are made solely through automated decision-making. One way the EDPB guidelines clarify the scope of Article 22 is by explaining on p.8 that: *"Solely automated decision-making is the ability to make decisions by technological means without human involvement"*. What human involvement entails is elaborated on p.21 of the EDPB Guidelines: *"To qualify as human involvement, the controller must ensure that any*

[17] ECLI:EU:C:2023:957, case C-634/21, Court of Justice of the European Union (2023).
[18] Supra note 12-16.

oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the relevant data".[19]

In Q7, the concept of 'human involvement' is linked to the extent to which the algorithm influences the outcome of the process, reflecting conceptual overlap with high-impact algorithms. By distinguishing whether the algorithm's output directly or determines the process outcome (option1-2), it becomes possible to assess whether human intervention is meaningful. Further details are provided in Q7.

Although profiling, as defined in article 4(4) of the GDPR, is referenced in the legal text of article 22(1), and frequently referred to in several of the consulted resources,[20] it has no bearing on the the scope of the provision and imparts no clarifying condition for qualifying sADM.

More information on the role of work instructions for human decision makers and the relationship between automated-decision making and the obligation to prevent discrimination can be found in the paper *"Meaningful human intervention for risk profiling algorithms – Preventing decision-making based solely on profiling".*[21]

Identification Q1 – What type of output does the application derive?

The type of output generated by an algorithmic system gives an indication of whether the system qualifies as sADM. See Figure 7.

When the output is a prediction (incl. a score, ranking, label, object-, face- or voice recognition), recommendation, decision or content, these outputs are assumed to be a weighting factor in the decision-making process that utilises these outputs (see section 3.1). If one of these options is selected the user is brought to Q2.

[19] Supra note 12.
[20] Supra note 12 and 13.
[21] Supra note 18.

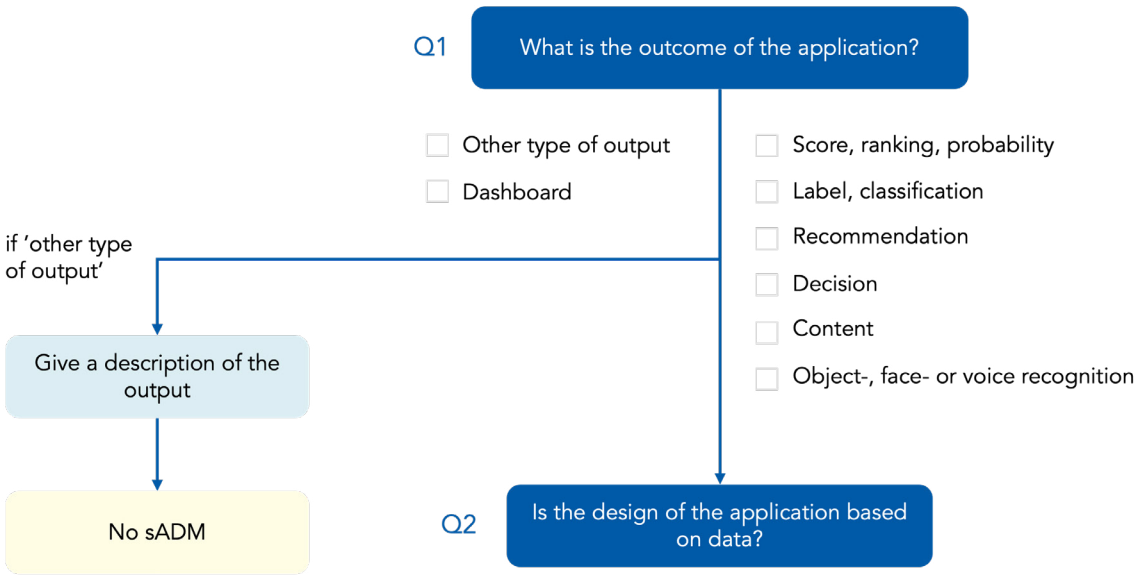


Figure 7 - Q1 of questionnaire Identification concludes that the application does not qualify as sADM if a 'Dashboard' or 'Other type of output' is selected.

Since a dashboard, on their own, only provides data visualization, there is no decision, legal or similar effects. It is the human that derives conclusions from this straightforward visualisation. If a user indicates the only output of a system is a dashboard, it is concluded it is not sADM. Users are explicitly prompted to consider if other types of outputs are displayed in this dashboard. If they select one of the specified outputs along with "Dashboard", they are brought to Q2.

The same logic applies to the option "Other type of output". If it is the sole provided answer, the tool concludes the application is not sADM. If coupled with another output (other than dashboard), the user proceeds to Q2. In either case, the user is asked to provide a description of the output, which can be manually assessed by experts.

Identification Q2 – Is the design of the application based on data?

Question Q2 provides no information for the qualification of sADM. The user answers this question because it supports qualification of AI and high-impact algorithms (see section 2 and 4).

Identification Q3 – Is the application an automation of rules defined by humans?

The only relevant difference in how Q3 captures the concepts of sADM and high-impact algorithms is that the form one-to-one automation cannot qualify as a high-impact algorithm but may still constitute sADM. More information about these cases are collected in questions Q8-Q12. In all cases users are brought to Q4 – in which information about the GDPR is collected. When the application is not an automation of rules defined by humans, users are requested to describe the process how rules are defined. See Figure 8.

Identification Q4 – Are personal data being processed by the application?

Since solely automated decision-making is a concept from the GDPR, Q4 inquiries whether personal data are being processed and whether the GDPR applies. If no personal data are processed, the process falls outside the scope of the GDPR and therefore cannot constitute sADM.

However, the absence of personal data does not mean an application is without significant impact. A neighbourhood profiling algorithm used to allocate

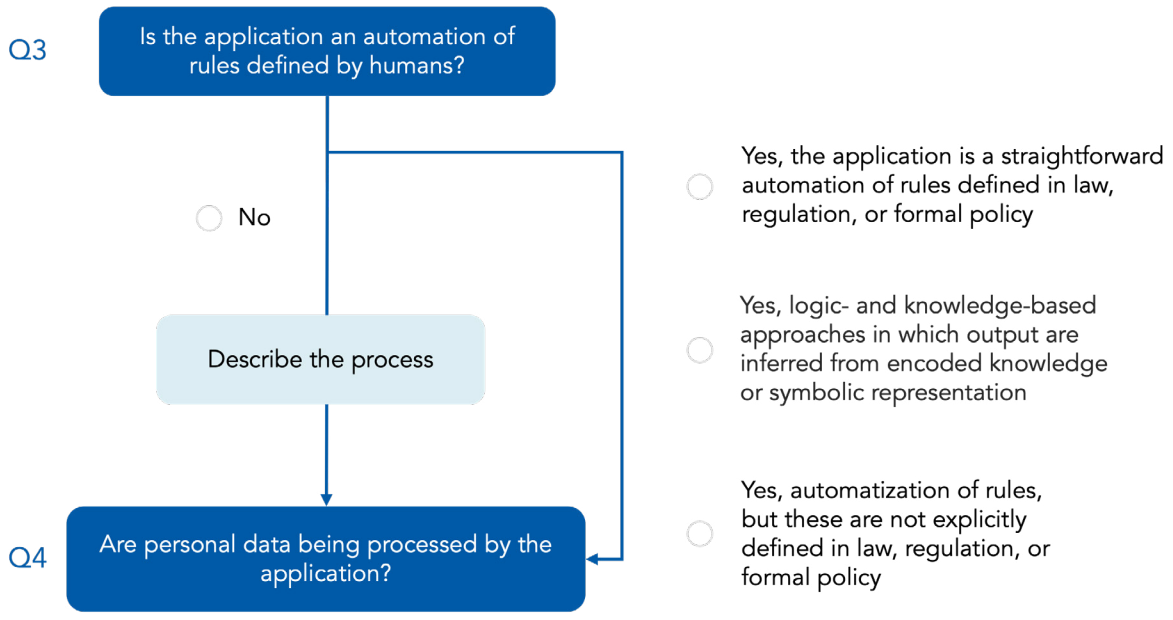


Figure 8 - All responses routes for Q3 in questionnaire Identification are directed to Q4 for GDPR identification.

police presence, for example, may have real consequences for individuals downstream. Yet, its output cannot be traced back to any specific person. Because no personal data is processed, there is no basis for prohibition under article 22 GDPR. Nevertheless, such a system may still qualify as a high-impact algorithm. These cases are discussed in section 4.

This distinction also matters in terms of scope: sADM applies only to decisions affecting individuals, whereas high-impact algorithms can affect groups as well (see section 4.1). For this reason, Q8 is included alongside Q4 to capture systems that may not trigger GDPR obligations but still carry significant societal impact. See Figure 9.

In case the output of the algorithm concerns personal data, the user is directed to Q5 to assess how the output of the algorithm is used.

If no personal data are processed, the user is still redirected to Q8. This follows the better safe than sorry principle of the design. In practice, users are expected to also answer Q8 with 'No.' Since Q8 can also capture whether a decision pertaining to an individual is made and therefore personal data should be processed. The user is not immediately presented the conclusion that sADM is not at hand when answering 'No' to Q4.

To assist users, the following remark is provided:

Examples of non-personal data are:

- Group statistics where no individual can be singled out
- Anonymised data
- Output about physical matters that are not linked to an individual, such as sensor readings, weather data, machine and operational data, urban planning and infrastructure data
- Financial records about companies
- Data about city districts and neighborhoods

NOTE: businesses in which the entrepreneur is personally liable (self-employed, sole proprietorship, general partnership, professional partnership) are considered as personal data.

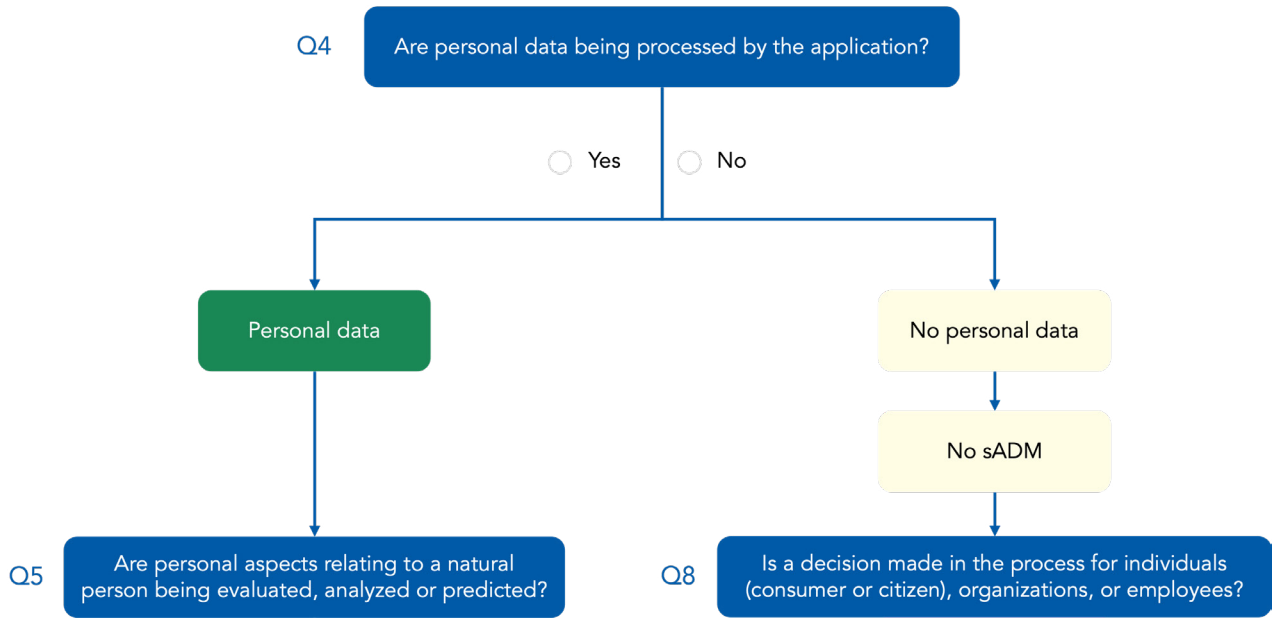


Figure 9 - Q4 of questionnaire Identification determines whether personal data are processed or not.

Identification Q5 – Are personal aspects relating to a natural person being evaluated, analyzed or predicted?

To determine whether the application involves profiling within the meaning of Article 4(4) GDPR, users are asked a question drawn directly from the legal definition. See Figure 10. In all cases, the user is forwarded to Q6.

The answer to this question is linked to Q32 of the Risk category questionnaire, in which the presence of profiling in the context of high-risk applications of AI systems is questioned. When the user has answered 'Yes' to Q5 of the Identification questionnaire and 'No' to Q32 of the Risk category questionnaire, the answers contradict each other. The user is then prompted with the following warning: "According to questionnaire Identification, we expect this answer to be 'Yes'. Please reconsider carefully whether the AI system is used to automatically evaluate or predict characteristics, behaviour or personal aspects."

Moreover, Q33 of the Risk category questionnaire identifies whether the filter of Article 6(3) of the AI Act can be applied. When this filter is applied the AI system is classified as high-risk-with-exception. To apply the filter an assessment should be made by the user that explains that the AI system does not perform profiling. Since the user already made this assessment by answering Q5, this answer is reused in the assessment underlying Q33.

To assist users, the following remark is provided:

Personal aspects include, for example, analysing or predicting job performance, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.

Identification Q6 – Is the output of the algorithm shared with other organizations?

As follows from the Schufa ruling and consequent jurisprudence, sharing the output of algorithms may result in sADM (see section 3.1).[22] Q6 therefore assesses whether the output of the algorithm is shared with other organizations. See Figure 11.

When the output of the algorithm is shared with other organizations, users are presented at the end of the questionnaire a warning that sharing outputs may lead to prohibited sADM. This should be assessed with support of legal professionals.

[22] Additional explanation of the Shufa ruling can be found in supra note 17.

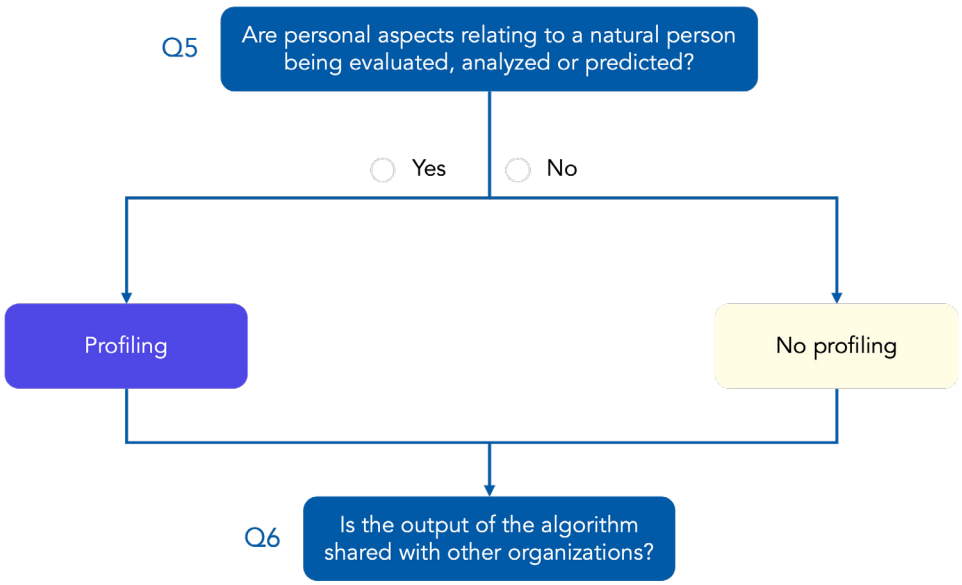


Figure 10 - Q5 of questionnaire Identification determines whether profiling, as defined in article 4(4) GDPR, is applicable.

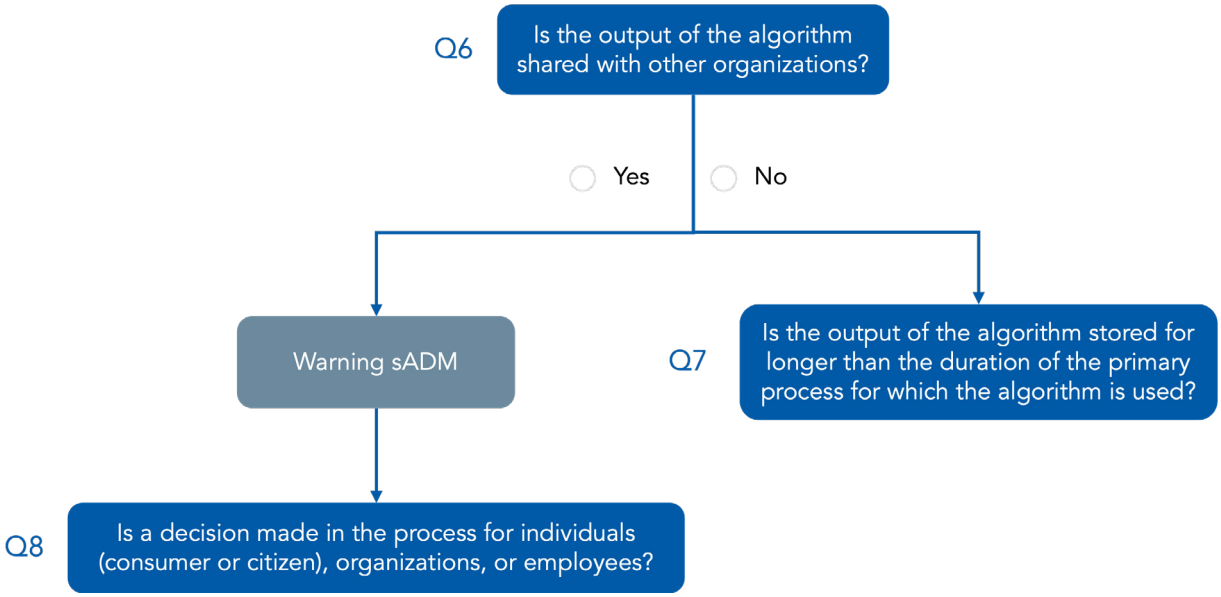


Figure 11 - Q6 of questionnaire Identification assesses whether the output of the algorithm shared with other organizations.

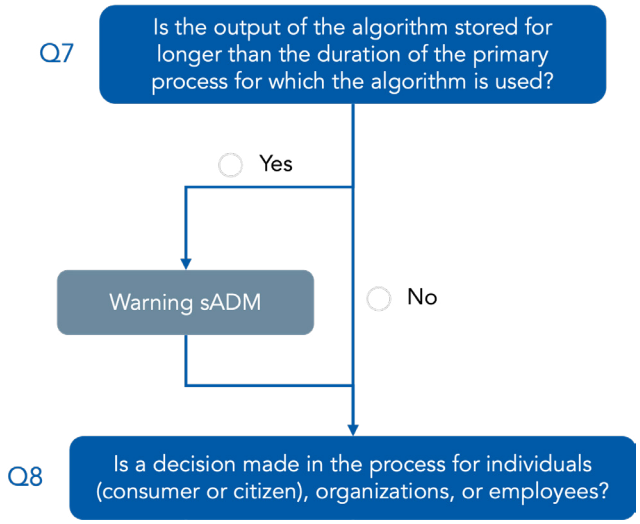


Figure 12 - Q6 of questionnaire Identification assesses whether the output of the algorithm shared with other organizations.

In case the output of the algorithm is not shared with other organizations, the user is directed to Q7.

Identification Q7 – Is the output of the algorithm stored for longer than the duration of the primary process for which the algorithm is used?

As follows from the Schufa ruling and consequent jurisprudence, storing the output of an algorithm may result in sADM (see section 3.1). Q7 therefore examines if the output of the algorithm is stored for longer than the duration of the primary process for which the algorithm is used. See Figure 12.

When the output of the algorithm is shared with other organizations, users are presented with a warning at the end of the questionnaire that sharing outputs may lead to prohibited sADM. This should be assessed with support of legal professionals.

In case the output of the algorithm is not shared with other organizations, the user is also directed to Q8 to continue the questionnaire.

Questions 8 to 12 also relate to the conditions that should be met for the filter of Article 6(3) of the AI Act to be applied, in which case an exception for the high-risk classification could be made. The answers to these questions are therefore reused in Q33 of the Risk category questionnaire.

Identification Q8 – Is a decision made in the process for individuals (consumer or citizen), organisations or employees?

Q8 assesses whether a decision is made in the process the algorithm is involved in. This is a necessary condition for sADM to apply (see section 3.1), though not a requirement for high-impact algorithms (see section 4). If the answer is 'Yes', users proceed to Q9. If the answer is 'No', users are redirected to Q11. This routing is designed to streamline the identification of high-impact algorithms. See Figure 13.

To assist users, the following remark is provided:

For instance:

> Prioritising citizen queries or requests
> Determining whether additional information is needed from a citizen
> Selecting individuals for checks or inspections
> Assessing eligibility for services or facilities.
> For public sector organisations, the following note is also included:

NOTE: A decision is much broader than a formal decision as defined in Dutch Public Administration Law. The notion of 'decision' is also used in the context of Article 22 GDPR.

Identification Q9 – What kind of decision is made in this process?

To identify sADM, it's also needed to assess what type of decision is made in the process the algorithm is involved in. This is assessed in Q9. This serves to assess whether direct consequences follow from the decision (see section 3.1). For sADM identification, a separate answer option is included to identify secondary effect of a decision made in the process in which the application is used ('Inspection, investigation or requests for additional information'), which is further discussed in Q10. See Figure 14.

Note that 'Prioritization of applications, requests, complaints, and objections' is not considered to be sADM, because this only concerns routing of tasks which doesn't affect the decision-making process itself. This differs from the logic used for the high-impact algorithm questionnaire, as the Algorithm

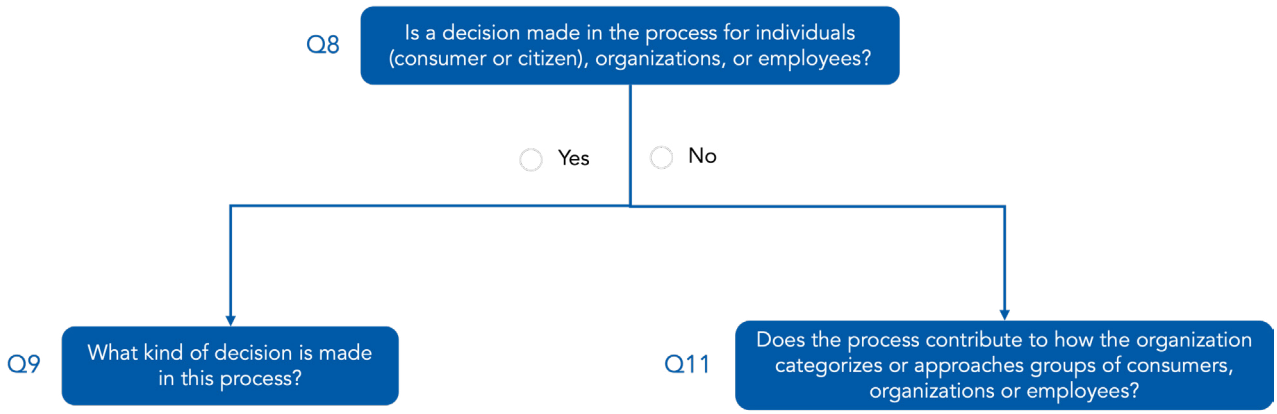


Figure 13 - Q8 of questionnaire Identification, determines whether a decision is made in the process the algorithmic application is used in.

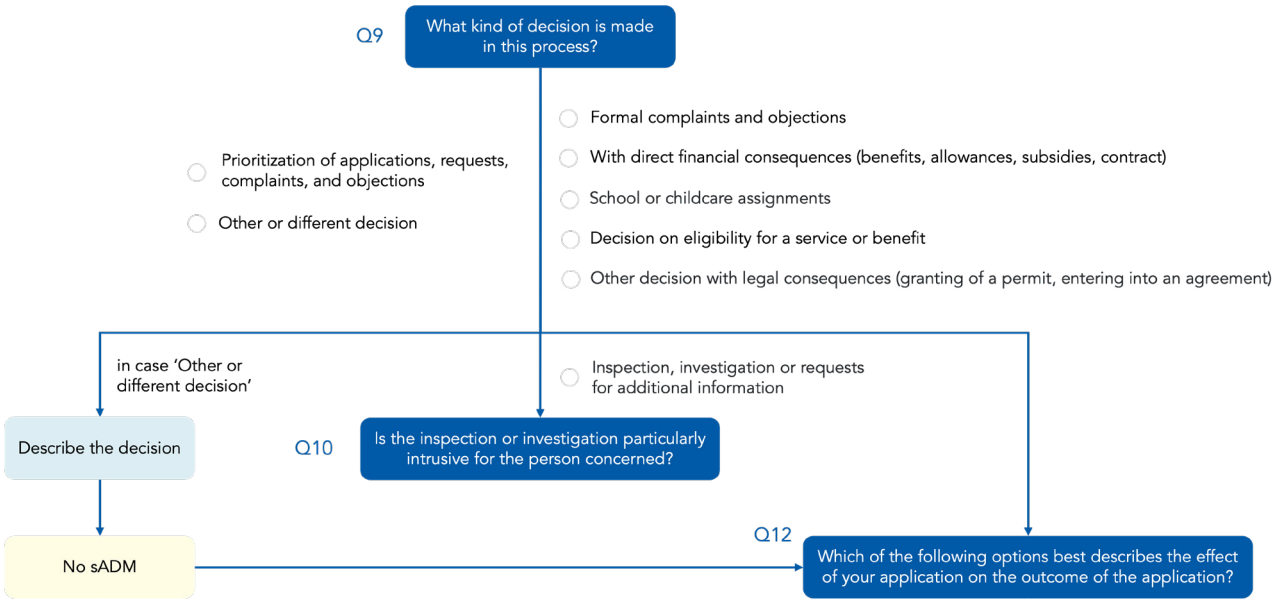


Figure 14 - Q9 of the questionnaire Identification, determines what type of decision is made in the process the algorithmic application is used in.

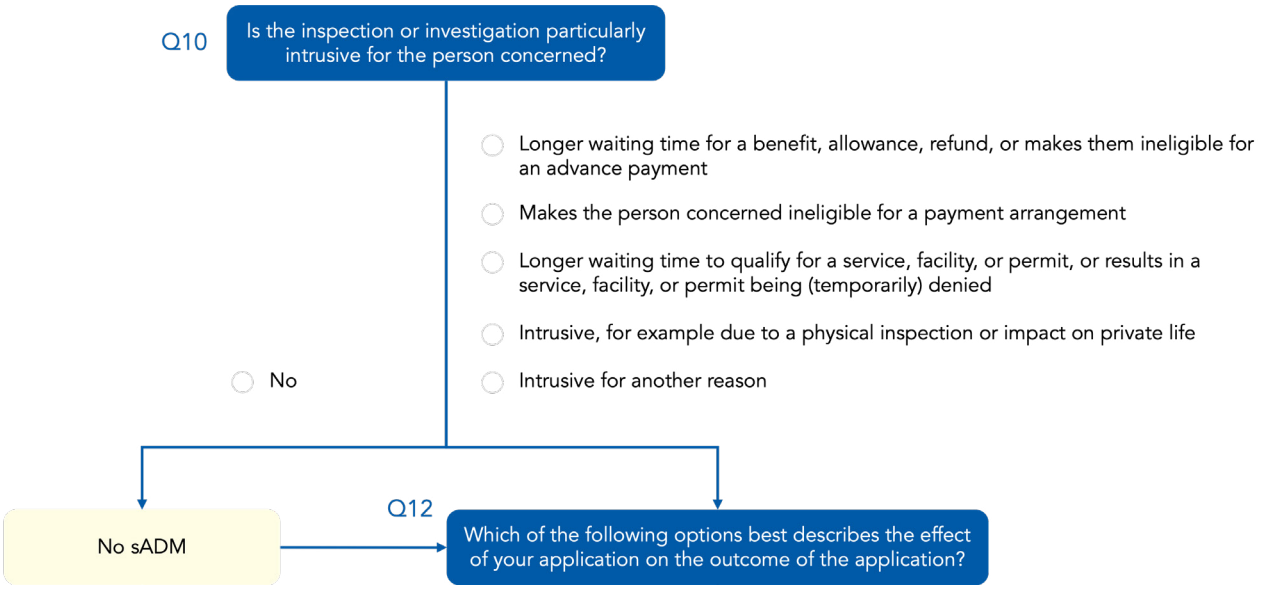


Figure 15 - Q10 of questionnaire Identification examines whether the inspection or investigation is particularly intrusive for the person concerned.

Registry Guidelines specifically mentions this falls in scope of the guidelines.

When 'Other or different decision' is selected, it can be concluded that no sADM occurs. The user is asked for a clarification. This can be reviewed by legal experts.

For all other answers, users are forwarded to Q12.

Identification Q10 – Is the inspection or investigation particularly intrusive for the person concerned?

As mentioned, Q10 is shown to the user only if 'Inspection, investigation, or requests for additional information' is selected in Q9. In this context, procedural decisions that do not have direct legal effects can still be considered to have a significant impact thourgh secondary effects (see section 3.1). This is assessed in Q10. See Figure 15.

When being selected for an inspection or investigation results in longer waiting time or affects eligibility for (unrelated) benefits or services, the user is redirected to Q12. The same applies if the inspection or investigation makes individuals ineligible for advance payments or payment plans, or significantly affects their private life – for example, through a home visit or other highly intrusive measures.

In case the inspection or investigation is not particularly intrusive for the person concerned, sADM is not applicable. Users are still redirected to Q12 to streamline the identification of high-impact algorithms (see section 4).

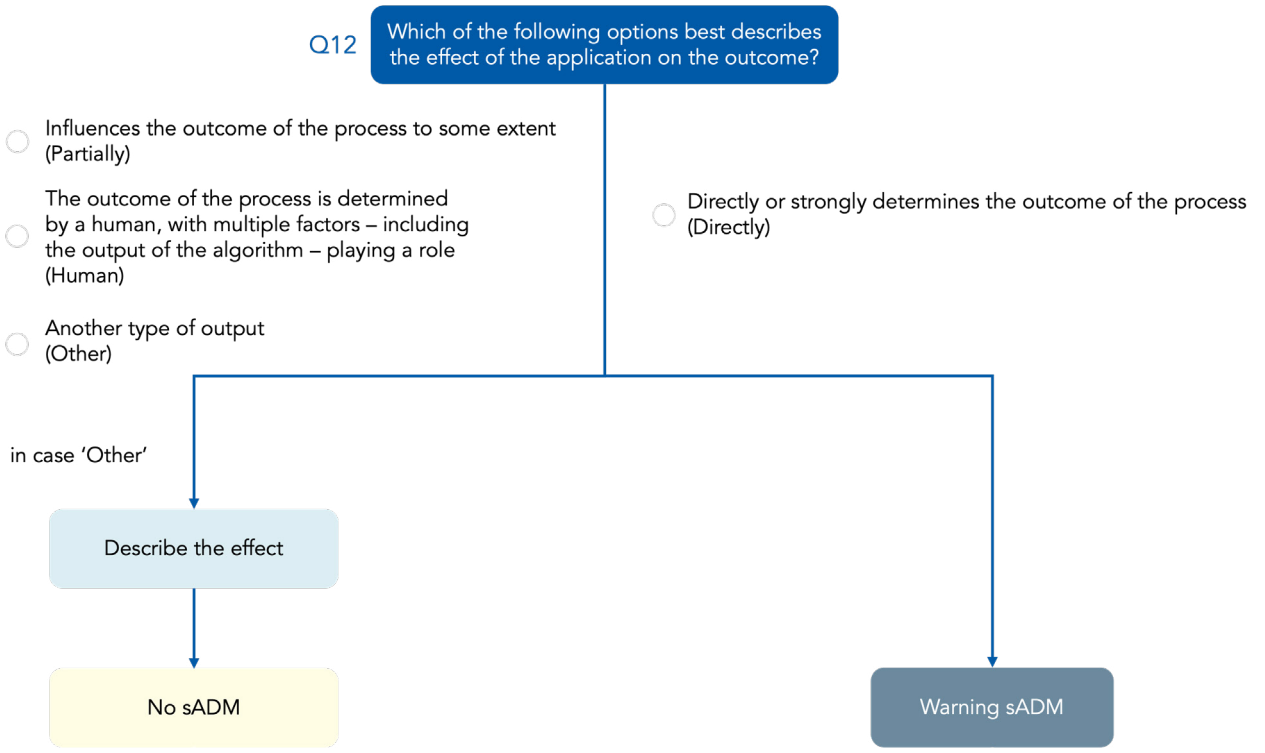


Figure 16 - Q15 of questionnaire Identification examines whether human intervention in the decision-making process is meaningful.

Identification **Q11 – Does the process contribute to how the organization categorizes or approaches groups of consumers, organizations or employees?**

For public sector organisations, the question is: "Does the process contribute to how the government categorizes or approaches groups of citizens, organizations, or civil servants?"

As Q11 is only shown when Q8 is answered 'No' (no decision is made in the process for an individual (consumer or citizen), organisation or employee), sADM is not at stake. Q11 is included to streamline the identification of high-impact algorithms (see section 4).

Identification **Q12 – Which of the following options best describes the effect of the application on the outcome?**

In the context of sADM, Q12 assesses the role of human involvement in the decision-making process and helps determine whether sADM is at hand. See Figure 16.

In the algorithmic system directly determines or largely influences the outcome of the process, sADM is at play. In all other cases, it is concluded that sADM is not at play due to human involvement. These conclusions are shown to the user. When 'Another type of output' is selected, the user is first asked to describe the effect.

Flowchart identification solely automated decision-making (sADM)

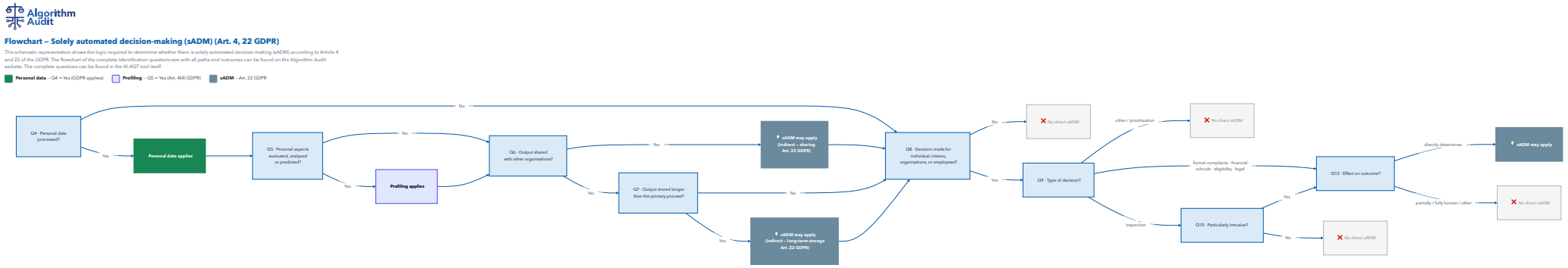


Figure 17 - Flowchart determining whether sADM is applicable or not.

4. Questionnaire Identification – High-impact algorithms

Algorithmic systems can have a significant impact on data subjects even when they do not qualify as an AI system under the AI Act (see Box 2). The term algorithm has been used since the 2010s – among others by the Dutch government – to refer to a broad category of automated systems, which also includes AI systems. In 2021, the Netherlands Court of Audit (2021) defined an algorithm as: *"A set of rules and instructions that a computer automatically follows when making calculations to solve a problem or answer a question"*.[23] An algorithmic system is a digital system through which an algorithm is executed. The definition of a high-impact algorithm, as set out in the "Algorithm Register Guidelines" issued by the Dutch Ministry of the Interior, is as follows:[24]

> **Direct consequences:** The algorithm has direct consequences for those involved (citizen, organization), e.g., imposing a fine or refusing a subsidy; or
> **Classification:** The algorithm influences how the government categorizes or approaches a data subject or group, e.g., profiling or risk indication for control.

The above categories are explained in the Algorithm Register Guidelines on the basis of three questions. See Figure 18.[25]

The three questions are the following:

1. Does it concern a process with direct consequences?
2. Are one or more algorithms used in the process?
3. Does the algorithm have a significant effect on the outcome of the process?

[23] 'Aandacht voor algoritmes', The Netherlands Court of Auditors (2021).
[24] Algorithm Register Guideline of the Ministry of the Interior and Kingdom Relations (2025).
[25] Style adapted from Algorithm Register Guideline of the Ministry of the Interior and Kingdom Relations (2025).

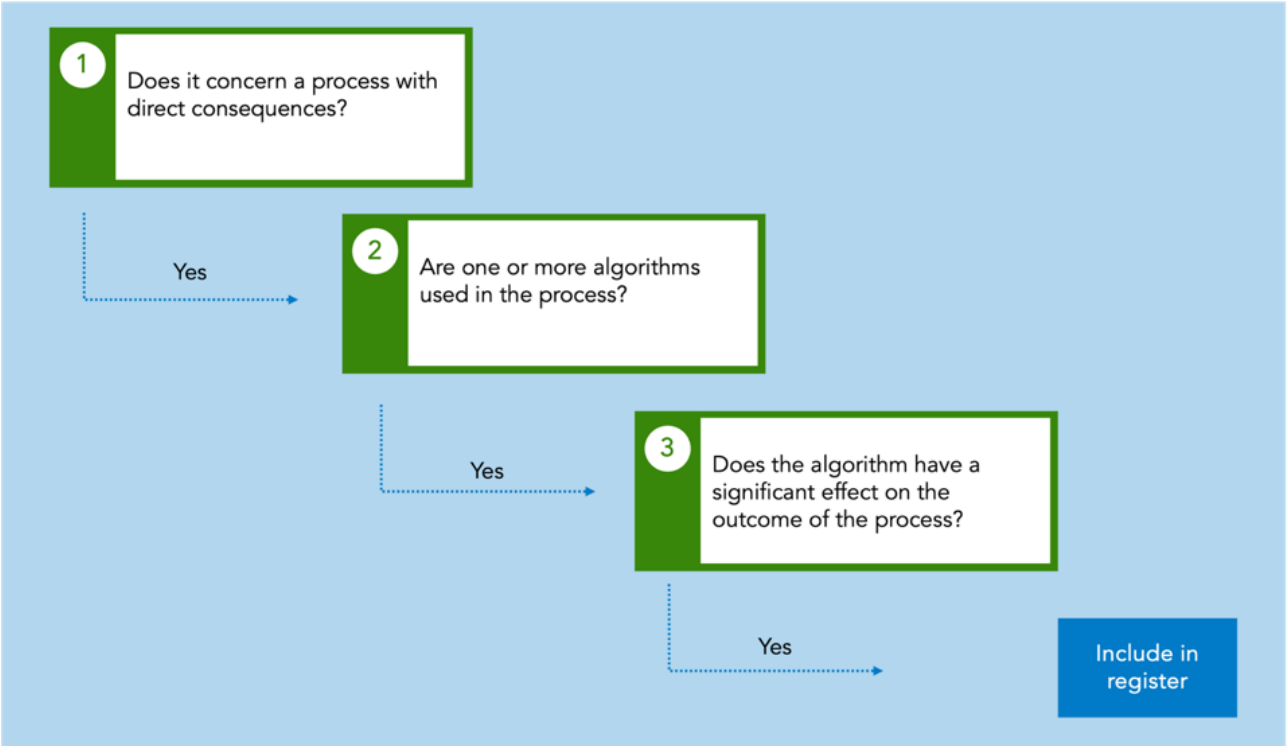


Figure 18 - Questions from Algorithm Register Guidelines that can be used to determine whether an algorithm is a 'high-impact algorithm'.

This section discusses how the first and last questions in Figure 18 – about the process in which the algorithm is used – are incorporated into the AI AQT as Questions 8-12 (Q8-12). Assuming at least one algorithm is involved in a given decision-making process, it can be deemed the second question in Figure 18 imparts no capacity to distinguish high-impact algorithms from other algorithms. The number of algorithms involved in a process is therefore not operationalised as a separate question in AI AQT. This section begins with an explanation of what is meant by the concepts of 'direct consequences' (section 4.1) and a 'significant effect' (section 4.2).

NOTE: For Dutch public sector organisations holds that high-impact algorithms must be published in the Dutch national Algorithm Register,[26] unless there is a ground for exception.[27]

[26] https://algoritmes.overheid.nl/nl/algoritme

[27] Grounds for exception as stated in the Guideline are: "legal grounds for exception as specified in the Open Government Act (Woo) and the Police Data Act (Wpg), or 'gaming the system'." See supra note 12.

Box 2

Dutch scandals were not caused by AI systems

Algorithms involved in scandals in the Netherlands, such as the childcare benefit scandal and the use of a discriminatory profiling algorithm by the Dutch Executive Agency for Education (DUO), involve algorithmic systems that fall outside the scope of the AI Act. These processes were based on human-defined rules without inference. In the case of DUO, students were assigned a risk score by a rule-based algorithm solely designed by human experts.[28] Although this algorithm does not qualify as an AI system, it had severe consequences: the risk profiling algorithm indirectly discriminated against students with a non-European migration background.[29] Likewise, the algorithm at the crux of the Dutch childcare benefit scandal also does not qualify as an AI system, yet it discriminated against Dutch citizens with dual nationality based on human defined decision rules.

High-impact algorithms appear to be more prevalent than AI systems, particularly in the Dutch public sector. In the summer of 2025, Algorithm Audit analysed the inventory of 14 Dutch ministries and concluded that 250 out of 370 (~67.6%) algorithmic systems qualify as high-impact algorithms, while only 13 out of 370 (~3.5%) qualify as high-risk AI systems.[30] This exemplifies that it is important to identify not only AI systems, but also high-impact algorithms so that appropriate control measures can be applied.

[28] Preventing prejudice, Algorithm Audit (2024).

[29] Addendum Preventing prejudice, Algorithm Audit (2024).

[30] Inventory 14 Dutch Ministries Netherlands Algorithm Registry, Algorithm Audit (2025).

4.1 Direct consequences

A high-impact algorithm is used in a process that has direct consequences for those involved. As clarified in the Algorithm Register Guidelines:[31]

I. *"These are processes with impact, which will generally be decision-making processes. Or the process contributes to how the government categorizes or approaches a person or group, for example by using weighting factors or predictions. This can have consequences for the approach or treatment. Examples of the latter are risk assessments and algorithms for fraud detection."*

II. *"In any case, the consequences include legal consequences. A legal consequence means that the decision under the Dutch Public Administration Law (Algemene wet bestuursrecht) affects the legal rights of a data subject, a person's legal status or their rights under an agreement. It also concerns factual consequences that affect the interests of a person, such as financial consequences (whether or not to receive an allowance), consequences for fundamental rights (whether or not to provide legal protection) and legal consequences (whether or not to stay in the Netherlands, to be allocated a home). The selection for an inspection or control is also seen as a consequence."*

III. *"Stakeholders include everyone who has to deal with the Dutch government. We summarize this as citizens and organizations."*

Direct consequences are broadly defined, which means many systems may qualify as high-impact algorithms. To determine whether an algorithm-driven process involves direct consequences, the first step is to assess whether a decision is made (Q8). For Dutch public sector organisations, the notion of a decision here is to be interpreted broadly: not only formal decisions, as defined in Dutch public administration Law (Awb Article1:3), may affect citizens and organisations. Other types of decisions may also have significant consequences for those involved and can therefore indicate the presence of a high-impact algorithm.

Once it has been established whether the process involves a decision affecting individual citizens or civil servants, the nature of the decision must be examined. This may include, for example, prioritisation of cases, decisions on formal complaints or objections or decisions with financial consequences (Q9).

Even in cases where no decision results from the algorithm-driven process, the system may constitute a high-impact algorithm if it alters how the government handles or approaches data subjects. Whether such an effect is present is assessed in Question 11 (Q11).

The following section discusses how the extent to which the algorithm impacts the outcome of the process is used to distinguishing high-impact algorithms from other algorithms.

4.2 Significant effect on the outcome of the process

A high-impact algorithm has a significant effect on the outcome of the process. As clarified in the "Algorithm Register Guidelines":

I. *"This is not about processes in which the algorithm automates/digitizes a manual work instruction. Such as algorithms in which all parameters are legally fixed and the algorithm runs through a (complex) decision tree based solely on these parameters".*

II. *"This does concern processes in which the algorithm influences a decision. Such as algorithms in which a weighting factor is given that (partly) determines the next step in the*

[31] Supra note 15.

process. The weighting factors are filled in by the space or freedom that an administrative body is entitled to in carrying out its tasks."

As explained in I., automation of rules that are 'legally fixed' in law, regulation or formal policy falls outside the scope of a high-impact algorithm. This can be dubbed 'one-to-one automation.' In such cases, the algorithm has no influence on the process – performed manually, the outcome would be no different. However, there are cases of '0.8-to-1 automation', where there is room for an organization to interpret a provision itself and formulate it as a decision rule. Such systems may qualify as a high-impact algorithm since any arbitrary specifications embedded into the algorithm shape the process. How a rule-based algorithm is design is assessed in Q3. Additional information about the effect of the algorithm on the outcome of the process is derived from Q12.

NOTE: The logic for the flowchart for the answer options in Q3 are different when assessing high-impact algorithms than when assessing an AI system. See 2. Questionnaire: AI system.

Identification Q1 – What type of output does the application derive?

The type of output generated by an algorithmic system gives an indication of whether the system influences the outcome of a process. Therefore, Q1 is used to determine whether a system could be considered a high-impact algorithm. See Figure 19.

When the output is a prediction (incl. a score, ranking, label, object-, face- or voice recognition), recommendation, decision or content, these outputs are assumed to be a weighting factor in the process that utilises these outputs (see section 4.2.II). If one of these options is selected the user is brought to Q2.

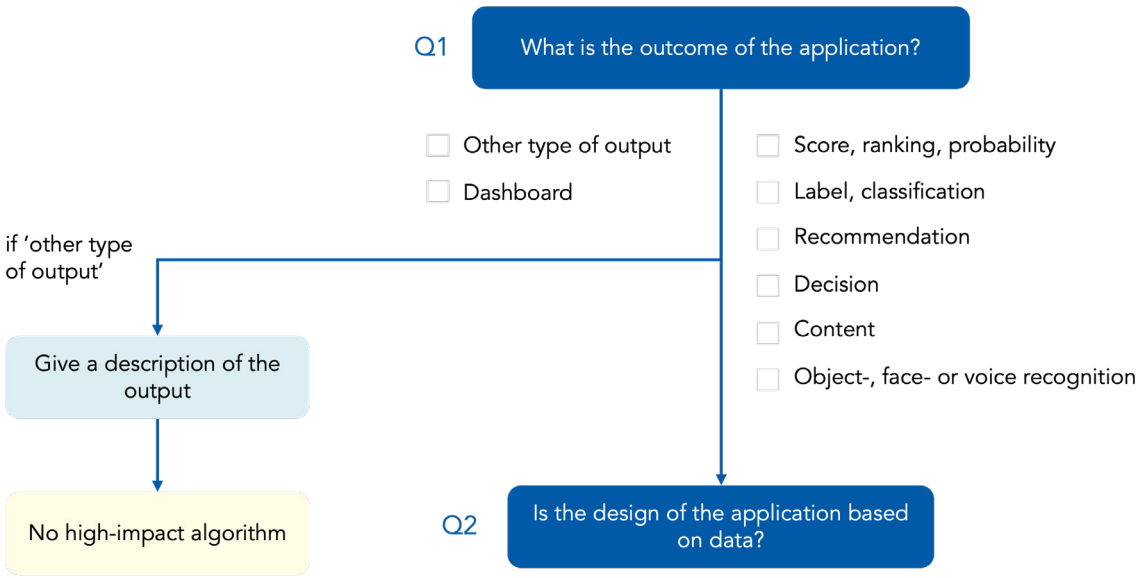


Figure 19 - Q1 of questionnaire Identification concludes that the application does not qualify as a high-impact algorithm if a 'Dashboard' or 'Other type of output' is at stake.

Since a dashboard, on their own, only provides data visualization, there is no direct consequence or significant effect at stake. It is the human that derives conclusions from this straightforward visualisation. If a user indicates the only output of a system is a dashboard, it is concluded it is not a high-impact algorithm. Users are explicitly prompted to consider if other types of outputs are displayed in this dashboard. If they select one of the specified outputs along with "Dashboard", they are brought to Q2.

The same logic applies to the option "Other type of output". If it is the sole provided answer, the tool concludes the application is not a high-impact algorithm. If coupled with another output (other than dashboard), the user proceeds to Q2. In either case, the user is asked to provide a description of the output, which can be manually assessed by experts.

Identification Q2 – Is the design of the application based on data?

If the algorithmic system is derived from data, it is not a straightforward automation of policy (see section 4.2.I). Therefore, Q2 supports qualification of high-impact algorithms.

If a system is derived from data or if the application contains components derived from data, the user is prompted to provide additional clarification and is sent to Q4.

When a user picks the option "no", they are sent to Q3 to understand whether this system should be excluded from the qualification of high-impact algorithm.

Identification Q3 – Is the application an automation of human-defined rules?

Whether the algorithmic system automates rules specified in law, regulation or formal policy or whether rules are defined by policy makers is an important indicator whether a system qualifies as a high-impact algorithm. See section 4.2 Significant effect on the outcome of the process and Figure 21.

In case the application is a straightforward automation of rules defined in law, regulation, or formal policy, the user must specify in which policy instrument the rule is stated. Thereafter, it is concluded that no high-impact algorithm is at stake (see section 4.2 Significant effect on the outcome of the process).

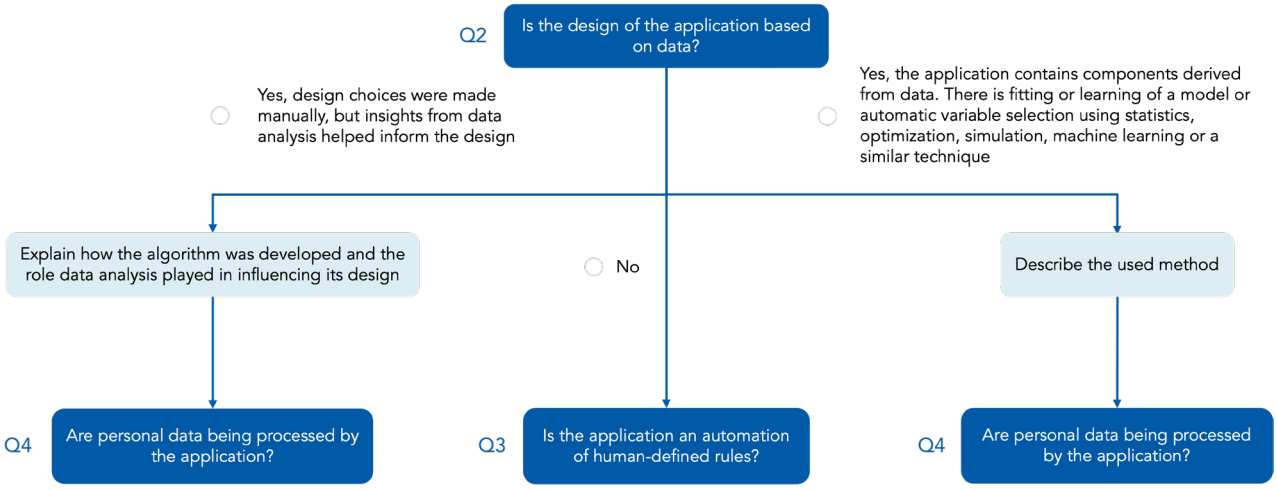


Figure 20 - Q2 of questionnaire Identification distinguishes whether more information is needed about the design of the algorithm (Q3), or that the user can continue to the next part of the questionnaire (Q4).

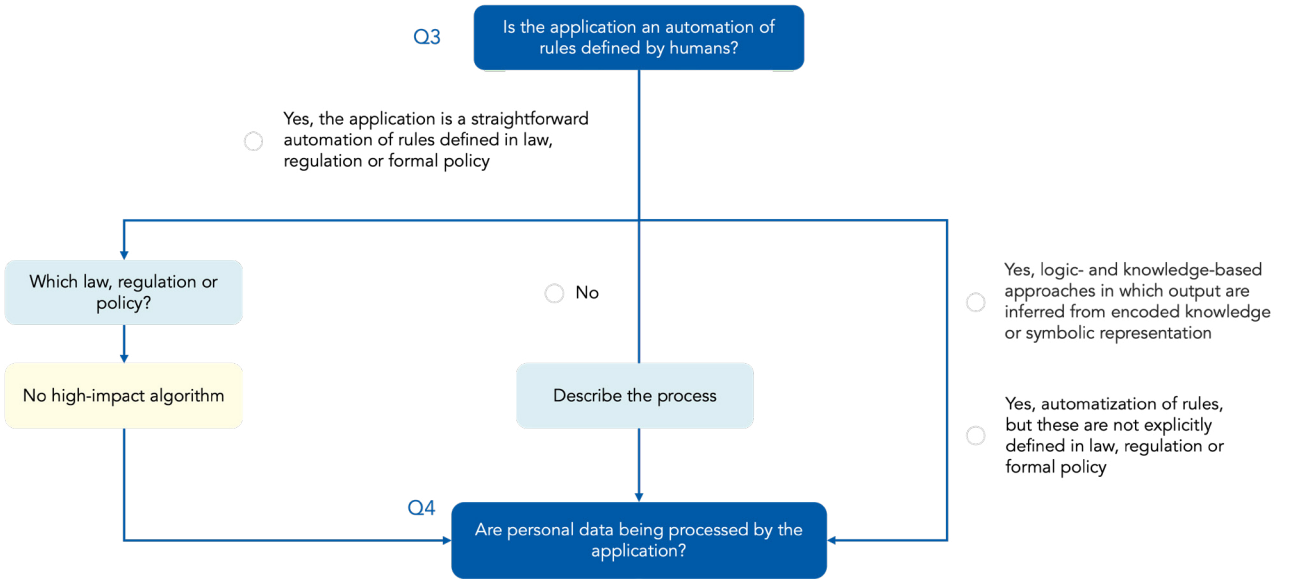


Figure 21 - In case of 1-to-1 automation of defined rules in law, regulation or formal policy, Q3 of questionnaire Identification concludes that the application is no high-impact algorithm. Other scenarios are forwarded to Q4.

If rules are not explicitly defined in law, regulation or formal policy, the user is forwarded straight away to Q4.

In case the application is not automation of human-defined rules, the user must describe the process and is then redirected to Q4.

An explanation for the answer option covering logic- and knowledge-based approaches can be found in 2. Questionnaire Identification: AI system, as this option is only relevant for AI systems.

NOTE: For this question, AI AQT recommends that work instructions for human decision-makers fall under answer option 'automatization of rules, but these are not explicitly defined in law, regulation or formal policy' rather than 'straightforward automation of rules defined in law, regulation or formal policy', as suggested in the Algorithm Registry Guidelines.

Identification Q4-Q7

Q4–Q7 don't play a role in identifying a high-impact algorithm. These questions are relevant for identifying sADM. A description of these questions can be found in section 3. Questionnaire: Solely automated decision-making.

Questions 8 to 12 also relate to the conditions that should be met for the filter of Article 6(3) of the AI Act to be applied, in which case an exception for the high-risk classification could be made. The answers to these questions are therefore reused in Q33 of the Risk category questionnaire.

Identification Q8 – Is in the process a decision made for individual citizens, organizations or employees?

To identify an essential characteristic of a high-impact algorithm (see section 4.1), it is first assessed whether a decision is made in the process the algorithm is involved. See Figure 22.

In case a decision is made, the user is forwarded to Q9. In case a decision is not made, the user is forwarded to Q11.

User guidance for this question is discussed in section 3. Q8.

Q4 relevant for qualifying high-impact and sADM

Q4 also supports the qualification of an algorithmic as sADM. This is further explained in 3. Questionnaire Identification – Solely automated decision-making.

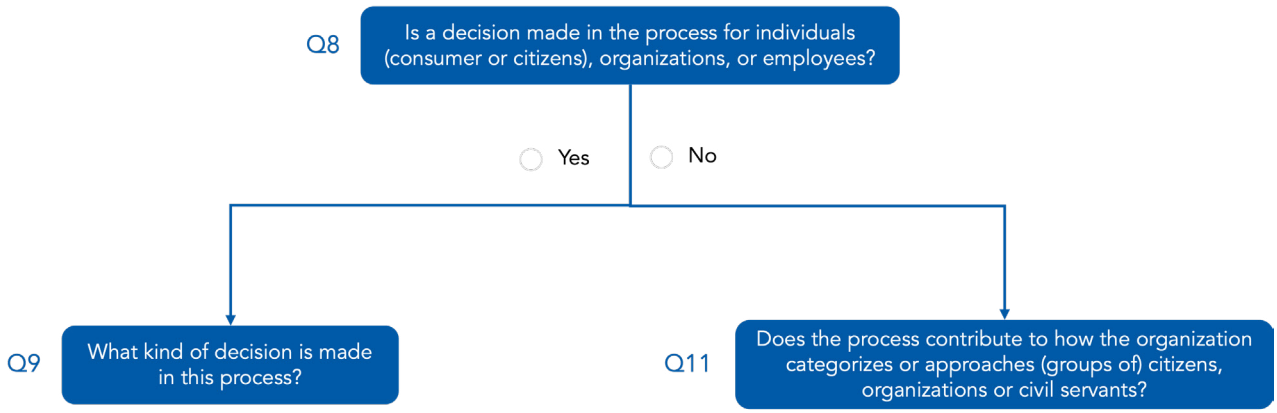


Figure 22 - Q8 of the questionnaire Identification examines whether a decision is made in the process in which an algorithm is involved.

Identification Q9 – What kind of decision is made in this process?

When a decision is made in the process the algorithm is involved in, it is important to assess the type of decision. See section 4.1 Direct consequencess and Figure 23.

A list of options is provided to help users identify the type of decision that can result in direct consequences for the data subject and may therefore indicate a high-impact algorithm. This list of options has been designed in collaboration with the municipality of Amsterdam, reflecting the breadth of decision-making with significant impact in the public sector. If one of these options is selected, the user is directed to Q12. Note that only one option can be selected. In future versions of AI AQT, answer options can be tailored to the specific sector in which the tool is used.

If a type of decision does not fall within any of the listed categories, it is concluded that the potential for direct consequences onto the subject is limited and that the algorithm in question does not qualify as a high-impact algorithm. In such cases, the user is requested to describe the type of decision, and the conclusion is shared with the user.

NOTE: Q9 also supports the qualification of a system as sADM. This is further explained in 3. Questionnaire: Solely automated decision-making.

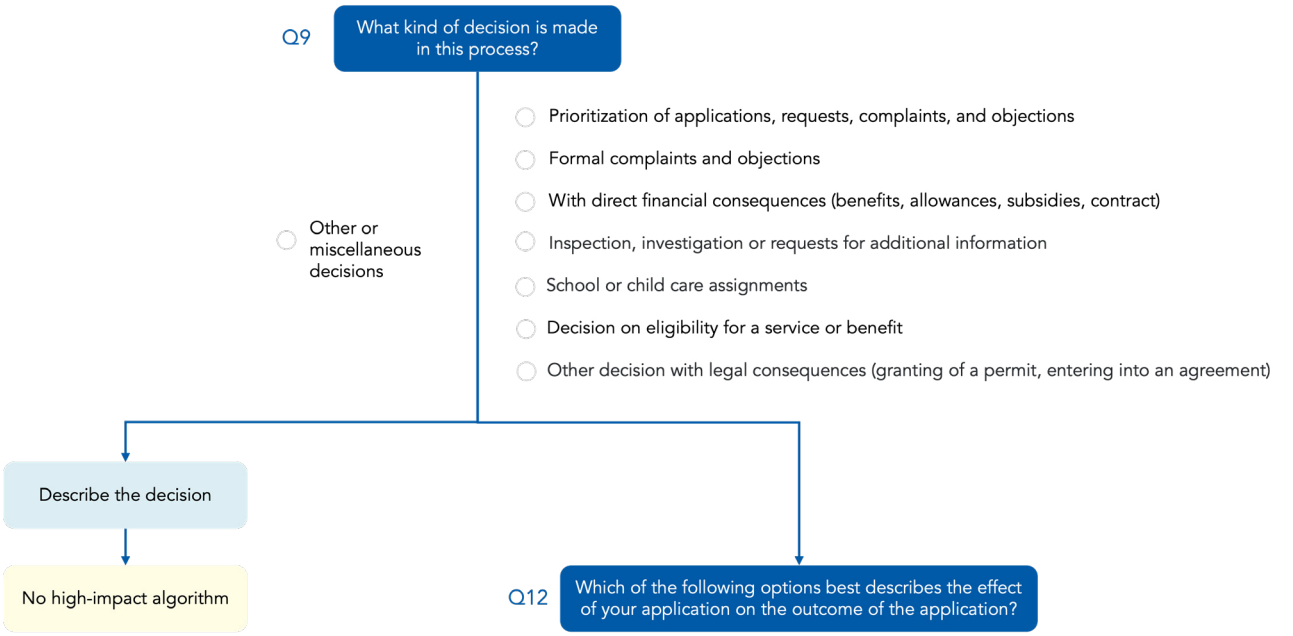


Figure 23 - Q5 of questionnaire Identification asks what kind of decision is made in the process in which the algorithm is involved.

Identification Q10 – Is the inspection or investigation particularly intrusive for the person concerned?

Q10 doesn't play a role in identifying a high-impact algorithm. This question is relevant for identifying sADM. A description of this question can be found in section 3. Questionnaire: Solely automated decision-making.

Identification Q11 – Does the process contribute to how the organization categorizes or approaches groups of consumers, organizations or employees?

For the public sector profile, the question is: "Does the process contribute to how the government categorizes or approaches (groups of) citizens or civil servants?".

Given that no decision is made in the process for an individual citizen, organisation or employee, as results from Q8, Q11 examines whether the process contributes to how the organization categorises or approaches groups of individuals, organisations or civil servants. See Figure 24. This is the second aspect of checking whether a direct consequence might follow from applying the algorithm (see section 4.1).

If the process does not contribute to how the government categorises or approaches groups of consumers, organisations or employees, it is concluded that the algorithm in question is not a high-impact algorithm and this conclusion is shared with the user.

If this cannot be said with certainty, an explanation is requested, after which the user is directed to Q12. The user will also be redirected to Q12 if Q11 is answered with 'Yes'.

NOTE: Q11 deliberately has a broader scope than individuals alone. For example, profiling neighborhoods to assign police surveillance capacity does not result in a decision about individuals, but it does influence how groups of citizens or consumers are approached by organizations or governmental actors.

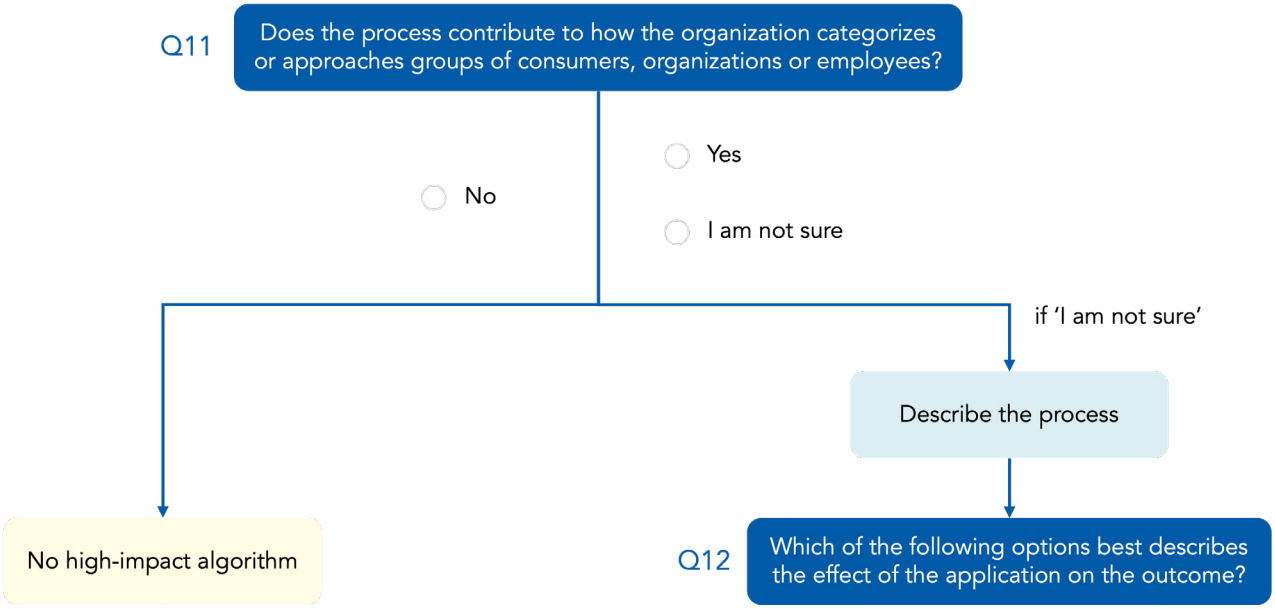


Figure 24 - Q11 of questionnaire Identification examines whether the process contributes to how groups of citizens, organisations or civil servants are categorised or approached by the government.

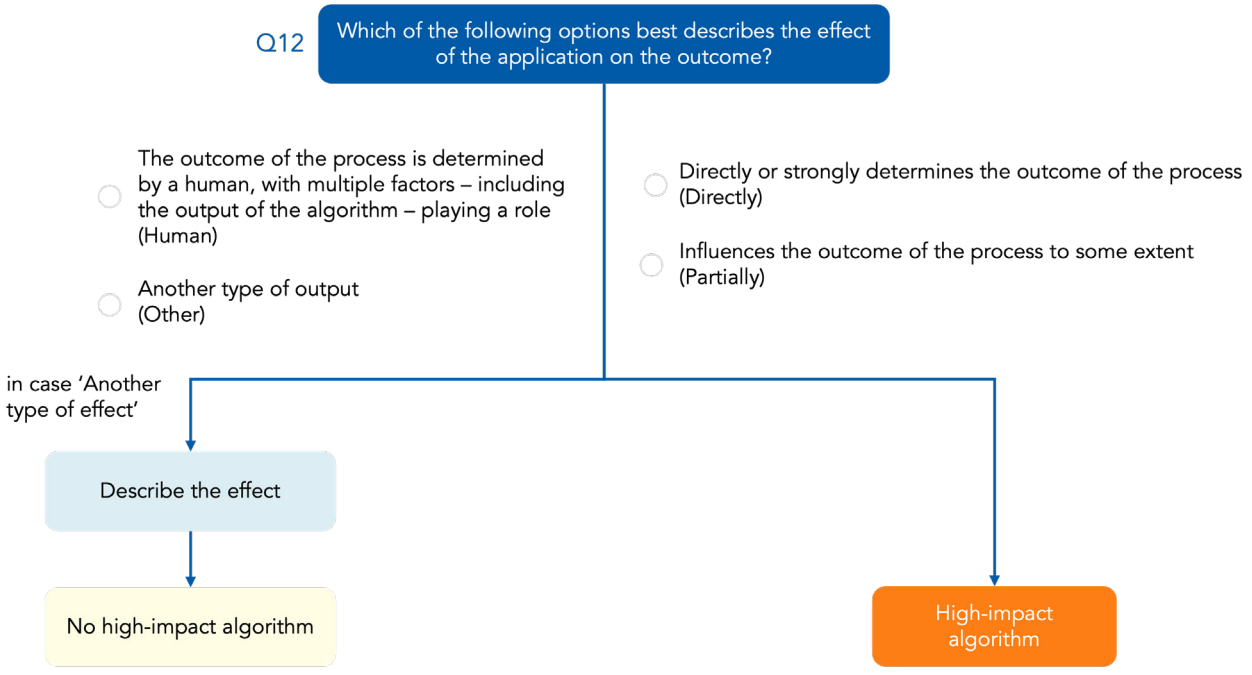


Figure 25 - Q12 of questionnaire Identification examines what the effect of the outcome of the algorithm is on the outcome of the process.

Identification **Q12 – Which of the following options best describes the effect of the application on the outcome?**

After it is established whether a decision is taken (Q8 and Q9) or the process alters how the organization approaches data subject (Q11),it must be determined whether the algorithm significantly effects the outcome of the process. If both conditions are met, it can be concluded that the algorithm is a high-impact algorithm. See Figure 25.

The scenarios described in the answer options assist users in selecting the most relevant description of the algorithm's effect on the outcome of the process. Where the outcome of the algorithm directly determines or partially influences the outcome to some extent, it is considered to have a significant effect and therefore qualifies as a high-impact algorithm.

Where the outcome of the process is determined by a human and multiple factors including the algorithm's output – play a role, the application is considered to not be a high-impact algorithm. The same conclusion applies where the application generates a type of effect other than those listed in the options. In such cases, users are asked to describe the effect so that the situation can be assessed manually by an expert team.

NOTE: Q12 also supports the qualification of a system as sADM. This is further explained in 3. Questionnaire: Solely automated decision-making.

Flowchart identification high-impact algorithm

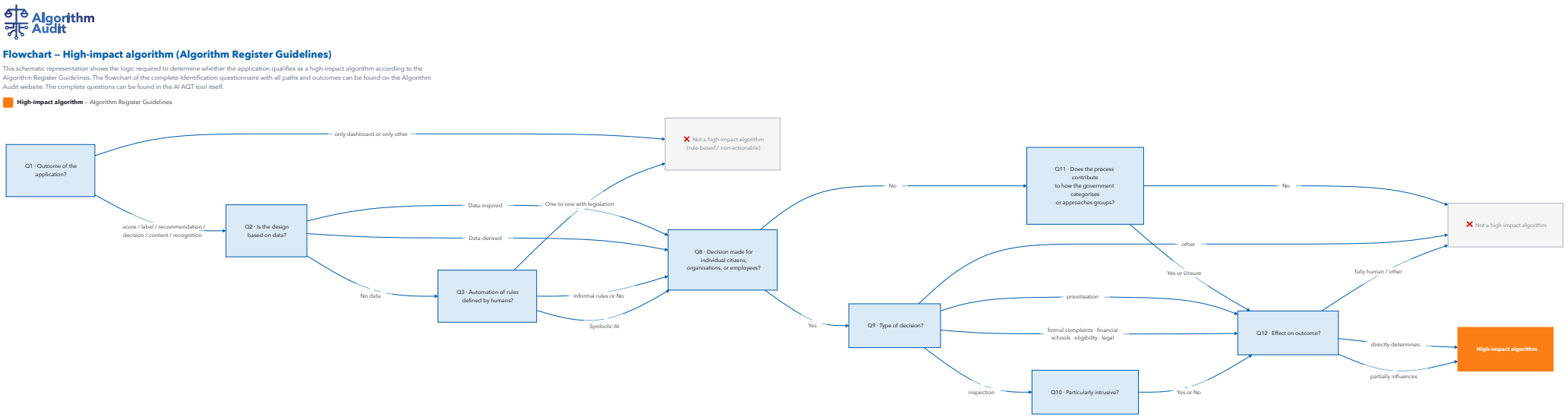


Figure 26 - Flowchart determining whether a high-impact algorithm is applicable or not.

5. Questionnaire
Role and status

The AI Act assigns obligations to specific actors involved in the lifecycle of an AI system. Articles 2 (Scope) and 3 (Definitions) of the AI Act describe these actors and the conditions under which the regulation applies to them. Once an algorithmic application has been identified as an AI system in Questionnaire Identification, the next step is to determine which role is held in relation to that AI system, since the role determines which obligations apply (Articles 16, 23, 24 and 26 AI Act).

The AI Act distinguishes five operator roles (Article 3 AI Act): provider, deployer, authorised representative, importer and distributor. A natural person who uses an AI system in the course of a purely personal, non-professional activity falls outside the scope of the AI Act (Article 2(10) AI Act). In addition to the role, the status of the AI system (whether it is already in use or still in development) determines the timelines for compliance (Articles 111 and 113 AI Act). AI systems that were already in use before the AI Act entered into force are subject to the transitional provisions of Article 111 and must comply with the requirements of the AI Act by 2030 at the latest. New AI systems face shorter, type-dependent deadlines under Article 113. Most notably, the deadline for most high-risk applications is 2 December 2027, while the prohibition on banned AI practices has applied since 2 February 2025.

This section first explains the operator roles (section 5.1) and the different usage status of an AI system (section 5.2). It then describes how these are operationalised through three questions (Q1, Q1.1 and Q2), which together make up the Role and status questionnaire.

5.1 Operator roles

It's the goal of questionnaire Role and status to assign one of the following operator roles, as stated in Article 3 of the AI Act:

- **Provider** – Natural or legal person that develops an AI system or general-purpose AI model and places it on the market under its own name or trademark (Article 3(3) AI Act).
- **Deployer** – Natural or legal person using an AI system under its own authority in a professional context (Article 3(4) AI Act).
- **Authorized representative** – A natural or legal person located or established in the Union who has received and accepted a written mandate from a provider of an AI system to carry out on its behalf the obligations and procedures established by the AI Act (Article 3(5) AI Act).
- **Importer** – Natural or legal person established in the Union who places on the market an AI system bearing the name or trademark of a non-EU entity (Article 3(6) AI Act).
- **Distributor** – Natural or legal person in the supply chain that makes an AI system available on the Union market without altering its properties (Article 3(7) AI Act).
- **Private user** – AI system is used in the course of a personal, non-professional activity (Article 3(4) AI Act).

Multiple roles can apply simultaneously. For example, an organisation that develops an AI system and uses it itself is at the same time provider and deployer. As Article 3(11) explains that "putting into service" is defined as the supply of an AI system for first use directly to the deployer or for own use in the Union for its intended purpose. So, an organisation that develops an AI system and starts using it internally is "putting it into service", which makes it a provider.

In this questionnaire, a 'Product manufacturer' is treated as a 'Provider'. A product manufacturer places an AI system on the market or puts it into service together with its product and under its own name or trademark is treated as a provider (Article 25(3)). Whether this exception applies depends on whether the AI system relates to products covered by EU harmonisation legislation listed in Annex I of the AI Act. Whether this is the case is determined in the questionnaire Risk category.

In this questionnaire, a 'Downstream provider' is also treated as a 'Provider'. A downstream provider is a provider that integrates a general-purpose AI (GPAI) model from a third party into its own AI system (Article 3(68) AI Act). Whether a GPAI model, in the form of a generative or interactive AI, is involved is addressed in questionnaire Risk category.

5.2 Status of the AI system

The AI Act draws a distinction between AI systems that are already in use and those that are still in development at the moment its obligations enter into force. This distinction is anchored in two provisions:

- **In use** – High-risk AI systems already placed on the market or put into service before 2 August 2027 must comply with the requirements of the AI Act by 2 August 2030 at the latest, unless the system is significantly changed in its design after that date (Article 111). Whether the high-risk classification applies is determined in questionnaire Risk category.
- **In development** – Read together with Recitals 177-179, sets out requirements for new AI systems, with type-dependent compliance deadlines (Article 113).

"In use" means that an AI system is actively being applied, including in a pilot or testing setting that affects work processes. Otherwise, the system is considered "in development".

Role and status **Q1 – How is the AI system developed or deployed?**

Q1 asks how the AI system is developed or deployed. The answer to this question is the primary determinant of the operator role. See Figure 27.

The answer options are mapped to operator roles as follows. Seven answer options assign one or two roles directly:

- "We only develop the AI system" → Provider.
- "We develop and use the AI system ourselves" → Provider and Deployer.
- "We develop the AI system and others also use it" → Provider and Deployer.
- "We import the AI system from a supplier outside the EU" → Importer
- "We buy the AI system from a supplier within the EU and resell it" → Distributor
- "We represent a non-EU person or entity to carry out AI Act obligations." → Authorised representative
- "You are a natural person using an AI system for purely personal non-professional activity." → Private user.

For all of the above answer options, after having answered Q1, users are redirected to Q3 to determine the status of the AI system.

The remaining answer option ("We use an externally developed AI system") routes to a follow-up question about responsibilities along the AI value chain (Q2). This question needs to be answered before an operator role is assigned.

Role and status **Q2 – Does one of the following scenarios apply to your AI system?**

Q2 determines if situations are applicable that a distributor, importer, deployer is considered to be a provider. This is the case when the following scenarios apply:

- Name and trademark: The name or trademark of your organization is put on the AI system (Article 25(1)(a)).
- **Substantial modification:** A substantial modification has been made to the AI system (Article 25(1)(b)). Beyond the formal legal definition in Article 3(23), the AI AQT describes a of a substantial modification as a significant change to an AI system's architecture, training data or functionality that materially alters how the system performs or the outputs it produces. When a high-risk AI system continues to learn after being placed on the market or put into service, this is not considered a substantial modification, provided the changes were pre-determined at the moment of the initial conformity assessment and described in the initial technical documentation (Article 43(4) AI Act).
- **Intended purpose:** The intended purpose of the AI system is modified by your organization (Article 25(1)(c)).

If one of the above scenarios apply, 'Provider' is being assigned as the operator role. If answer option 'None of the above' is selected, the 'Deployer' operator role is assigned. See Figure 27.

Whether a distributor, importer or deployer becomes a provider under the above scenarios mainly matters for high-risk AI systems. The Risk category questionnaire establishes the system's risk classification. In this questionnaire, it is determined whether an additional responsibility along the AI value chain applies.

Role and status **Q3 – Is the AI system already in use?**

Q3 determines the status of the AI system, which controls the compliance deadlines that apply to the user's role.

If the AI system is "In use", the user must comply with the requirements of the AI Act by 2030 at the latest (Article 111 AI Act). If the AI system is "In development", shorter, type-dependent deadlines apply (Article 113 AI Act, read in conjunction with Recitals 177-179).

After Q3 the questionnaire presents the result: the role(s) identified for the user, the status of the AI system. Responsibilities depend on the risk category of the AI system. It is therefore suggestion to users to fill the Risk category questionnaire.

Flowchart Role and status

Algorithm Audit

Flowchart – Role and status (Art. 2, 3, 25, 43, 111, 113 AI Act)

This schematic representation shows the logic required to determine what your role is in relation to the AI system and how the usage status of the system can be determined according to Article 2, 3, 25, 43, 111 and 113 of the AI Act. The complete questions can be found in the AI AQT tool itself.

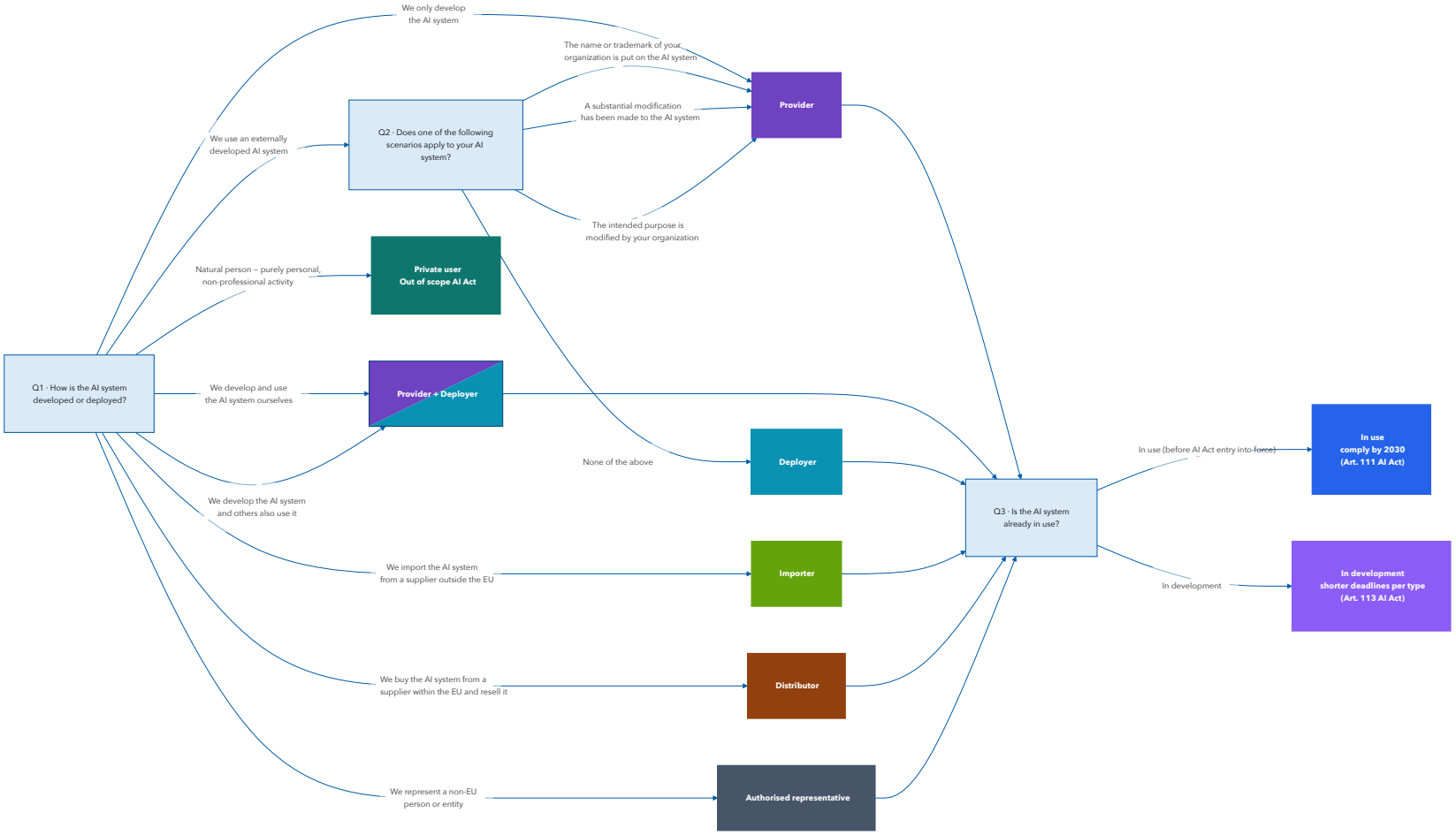


Figure 27 - Flowchart determining your role in relation to the AI system and status of the AI system.

6. Questionnaire Risk category

The AI Act of the European Union (EU) follows a risk-based approach: Depending on the risk category an AI system falls into different obligations apply (see section 7). In this section the risk category is determined. Questionnaire Risk category distinguishes five categories:

- **No requirements** – No obligations under the AI Act apply to the AI system (Articles 5 and 6). Obligations under other legislation (such as the GDPR) may still apply.
- **Prohibited AI systems** – The AI system meets the definition of a prohibited AI practice and may not be placed on the market, put into service or used within the Union (Article 5).
- **High-risk AI systems** – The AI system meets the classification rules for high-risk AI systems and must comply with the requirements depending on the role of your organisation in relation to the AI system. See section 7.
  Generative and interactive AI – Article 50 imposes specific transparency obligations on providers and deployers of generative or interactive AI systems. These obligations depend on the type of AI used. See section 7. These obligations apply regardless of the high-risk classification.
- **Exception** – The AI system would otherwise be high-risk or prohibited, but an exception for defence, military use or national security (Article 2(3)), scientific research and development (Article 2(6) and 2(8)) or narrow procedural/preparatory tasks (Article 6(3)) may apply. Verification with legal experts is required to confirm whether the exception is available.

Questionnaire Risk category determines, on the basis of 5-15 questions, into which of the above categories an AI system falls. The questionnaire opens with two gate questions that decide which branch is followed. First, it is checked whether Annex I – existing EU product-safety laws that are relevant for classifying AI systems as high-risk – is applicable (Q1). Then, it is checked whether the AI system uses biometric data (Q2). Both questions need to be answered before the questionnaire branches into more specific assessments.

The sections describe the questions in the order in which they are asked: questions about Annex I (6.1), the biometrics branch (6.2), transparency obligations and non-consensual fake nude imagery (6.3), nine high-risk domains from Annex III (6.4), horizontal Article 5 prohibitions (6.5), Article 6(3) profiling and limited-task exceptions (6.6) and Article 2 exceptions (6.7). This chapter concludes with the overall flowchart of the Risk category questionnaire in Figure 28.

The questionnaire is updated according to the 'Draft Commission guidelines on the classification of high-risk AI systems'.[32]

6.1 Annex I – List of Union harmonisation legislation

Risk category Q1 – Does your AI system fall within the scope of Annex I of the AI Act — either as the product itself, or as a safety component of a regulated product?

Article 6(1) classifies an AI system as high-risk if it is intended to be used as a safety component of a product covered by EU harmonisation legislation listed in Annex I of the AI Act, or if the AI system is itself such a product, provided that the product must undergo a third-party conformity assessment under that legislation. See also considerations 23 through 28 of the draft guidelines on the classification of high-risk AI systems of the AI Act. Section A of Annex I lists the modern EU product-safety regimes

[32] Draft Commission guidelines on the classification of high-risk AI systems under Article 6 of Regulation (EU) 2024/1689 (AI Act) for stakeholder consultation (2026).

built around the New Legislative Framework (NLF) – a standardised approach to CE marking, conformity assessment, accreditation, and market surveillance.[33]

Annex I – Section A:

- Toy Safety Directive (Directive 2009/48/EC)
  - Includes AI safety functions in connected toys. Does not include an AI system that recommends music in a toy.
- Recreational craft and personal watercraft (Directive 2013/53/EU)
  - Examples include AI controlling navigation, propulsion safety, collision avoidance, or stability on recreational craft and personal watercraft — systems whose failure or intended safety function affects passenger safety.
- Lifts and safety components of lifts (Directive 2014/33/EU)
  - Examples include AI managing door closing timing and obstacle detection. The provider's intended purpose may be efficient lift operation, but a malfunction could cause injury and thus qualifies as a safety component.
- Equipment for explosive atmospheres (Directive 2014/34/EU)
  - Such as AI monitoring gas concentrations in explosive-atmosphere equipment and commanding shutdown.
- Radio equipment (Directive 2014/53/EU)
  - Such as a smart thermostat where the intended purpose is a safety function (e.g. preventing carbon monoxide formation or fire risk) or where failure may endanger health or safety beyond mere discomfort. A smart thermostat that optimises comfort or energy efficiency only is not included.
- Pressure equipment (Directive 2014/68/EU)
  - Such as AI predicting runaway pressure in pressure equipment and actuating protective measures such as shutdown linked to safety accessories.
- Cableway installations (Regulation 2016/424)
  - Examples include AI supervising drive or brake systems, detecting collisions, or controlling stability on cableway installations.
- Personal protective equipment (Regulation 2016/425)
  - Examples include AI sensors integrated into personal protective equipment that detect hazardous conditions, monitor vital signs of the wearer, or trigger protective responses.
- Gaseous fuels (Regulation 2016/426)
  - Examples include AI optimising combustion efficiency in household gas appliances where failure could lead to carbon monoxide formation, explosion, or fire. AI that merely optimises heating schedules based on household habits where failure results only in discomfort or higher bills is not included.
- Medical devices (Regulation 2017/745)
  - Examples include AI diagnostic, monitoring, alerting, or treatment-decision-support systems integrated into regulated medical devices. Both stand-alone medical-device AI (where the AI system itself is the product) and embedded safety functions can qualify.
- In Vitro diagnostic medical devices (Regulation 2017/746)
  - Examples include AI systems analysing in vitro samples for diagnostic results, alerting on critical values, or supporting treatment decisions where misdiagnosis arising from system failure could endanger health.

Systems that relate to a Section A product are subject to the full set of high-risk obligations (see section 7). These obligations are integrated into the existing sectoral conformity-assessment procedures, so manufacturers typically go through a single combined assessment rather than two parallel ones.

[33] The 'Blue Guide' on the implementation of EU product rules, European Commission (2022).

Section B lists sector-specific regimes that fall outside the NLF, mainly covering transport and aviation. The AI Act imposes no direct obligations for providers or deployers of these systems. Their duties come from the operator-side rules of the sectoral framework.

Annex I – Section B:

> Machinery (Directive 2006/42/EC)
  - Examples include an AI based computer vision system that detects human presence in a robot cell and triggers a safe stop or speed reduction where the intended purpose is to prevent injury.
> Civil aviation security (Regulation 300/2008)
  - Examples include AI supporting passenger or baggage screening, threat detection, or access control to security-restricted zones in civil aviation.
> Two- and three-wheel vehicles and quadricycles (Regulation 168/2013)
  - Examples include AI for stability control, anti-lock braking, collision avoidance, or rider-assist systems on two- or three-wheel vehicles and quadricycles whose failure could endanger health or safety.
> Agricultural and forestry vehicles (Regulation 167/2013)
  - Examples include AI targeting agricultural land for chemical spraying that does not adequately account for nearby persons could endanger health and safety. AI used solely for yield forecasting or irrigation optimisation is not included.
> Marine equipment (Directive 2014/90/EU)
  - Examples include AI for navigation, collision avoidance, stability and ballast control, or lifesaving equipment — systems whose intended safety function or failure consequences affect vessel or passenger safety.
> Rail system interoperability (Directive (EU) 2016/797
  - Examples include AI in a train monitoring speed limits to prevent collisions and derailments.
> Motor vehicle type-approval (Regulation (EU) 2018/858 and Regulation (EU) 2019/2144)
  - Examples include AI lane assistance, AI for advanced driver-assistance, automated emergency braking, drowsiness or distraction detection, or other safety functions required under the General Safety Regulation.
> Civil aviation / EASA rules, insofar as they concern unmanned aircraft (Regulation (EU) 2018/1139).
  - Examples include AI in flight controls, collision avoidance (TCAS-type) systems, engine or system-health monitoring, or air-traffic management aids — systems whose failure could endanger flight safety.

To assist users, the following remark is provided:

> Choose 'yes' when the AI system is embedded in the product, or supplied independently as a software update, add-on or remote service. Consumer products are not excluded.
> The AI system is the product itself when it's independently placed on the market with its own intended purpose. Example: under the Machinery Directive, certain software is itself a regulated product.
> The AI system is a safety component of a regulated product, when it qualifies under either one of the two following options:
  i. Safety function: the intended purpose of the system, as documented in instructions of use, technical documentation or promotional materials, is to prevent or mitigate risks to health and safety of persons or property. Examples are: AI-based safe-stop trigger in a

robot cell, AI monitoring gas concentrations and commanding shutdown, AI predicting runaway pressure and actuating protective measures.

ii. Failure or malfunctioning endangering health and safety of persons or property: even where the provider does not intend a safety function, the system qualifies if its failure or malfunction (incorrect outputs, drift, latency, misclassification) would endanger health, safety or property. Examples: AI for lift door timing and obstacle detection, AI lane-assistance in a vehicle and AI for chemical-spraying drones.

> Examples that do not fall within the safety-component definition: AI that recommends music in a connected toy, AI drones used solely for forecasting the amount of crop that will be produced in a given area or irrigation optimisation.

When the user selects 'None of the below', the questionnaire continues to the biometric data screen (Q3). When the user selects an Annex I directive or regulation, the questionnaire presents Q2, which asks whether the product requires a third-party conformity assessment under that legislation.

Risk category **Q2 – Is the product whose safety component is the AI system, or the AI system itself as a product, required to undergo a third-party conformity assessment under EU harmonisation legislation?**

Under Article 6(1), an AI system is classified as high-risk through the "safety component" route only if the same Annex I legislation selected in Q1 also requires the product (or the AI system itself, where it is the product) to undergo a third-party conformity assessment. This is what Q2 asks. See also Recital 50 and 51 and considerations 50 through 59 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

If the answer to Q2 is 'Yes' in combination with Annex I Section A legislation in Q1, obligations for high-risk AI systems apply (see section 7).

If the answer is 'Yes' in combination with Annex I Section B legislation in Q1, the AI system is classified as high-risk but the AI Act's own high-risk obligations (Articles 8-27) do not apply. Only Article 6(1), Articles 102 to 109 and Article 112 apply. Articles 102 to 109 require the underlying sectoral product laws (aviation, vehicles, marine equipment, etc.) to embed AI Act specific requirements directly into each law's existing requirements and conformity assessment procedure. Providers thus carry the compliance responsibility through the sectoral framework that already governs their product. Article 112 imposes no obligations on providers of this type of AI system. Instead, it empowers the Commission to periodically review the AI Act and, where necessary, amend it in this respect.

When the answer is 'No', the user is forwarded to the biometric data screen (Q3), as the AI system can still be high-risk under Annex III.

To assist users, the following remark is provided:

If you already have a product that is subjected to existing product-safety compliance (e.g., B-type conformity assessment, obligatory notified body involvement). Check the conformity assessment route in your documentation. More information about different types of third-party conformity assessments can be found on the Algorithm Audit website.[34]

Risk category **Q3 – Does the AI system use biometric data?**

According to Article 6(2), AI systems referred to in Annex III are also considered to be high-risk. Annex III(1) lists AI systems that use biometric data as a high-risk category. Biometric data are defined in

[34] Infographic third-party conformity assessment, Algorithm Audit (2026).

Article 3(34). See also Recital 14 and considerations 122 through 125 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

If the AI system uses biometric data ('Yes'), the user is forwarded to the biometrics branch (Q4). If the answer is 'No', the user is forwarded to the Article 50 transparency scenarios (Q11).

To assist users, the following remark is provided:

Biometric data are unique characteristics of a person such as images or video footage of faces, iris scans, fingerprints, voice sounds, handwriting, and other unique characteristics of bodies and/or behaviour.

The key distinction of biometric data is: does the data measure bodily, structural and relatively static attributes of a person or does it merely describe or record something associated with a person?

Examples of biometric data:
- Fingerprints, iris or retina patterns
- Facial and hand geometry, voice print, vein patterns, ear shape
- DNA, odour, heartbeats, electroencephalography (EEG) and electrocardiograms (ECG)
- Someone's unique way of typing (keystrokes), walking (as captured by sensors or video analysis) or distinctive characteristics of movements, gestures, and motor-skills of natural persons.

Examples of non-biometric data:
- A photo of someone
- Typed text or written signature
- Name or ID number
- Medical diagnosis
- Location or GPS data
- Passwords
- IP address or device fingerprints (tied to devices, not bodies)
- Age or birth date, height and weight
- An AI system that follows a person based on, for example, a red scarf or a blue jacket (no biometric data)
- Data from deceased individuals, dead bodies, human remains or traces thereof (biometric data only concerns living natural persons)
- Object recognition. For instance, an AI system used in supermarkets to detect whether a product is being put into a basket, or suspicious objects detection at airports, train stations etc.
- Emotions inference from typed text (content, sentiment analyses or stylometry).

See also: Physiological and Behavioural Biometrics – Biometrics Institute

6.2 Biometrics branch

Risk category **Q4 – What does the AI system do with biometric data?**

Q4 splits the biometrics branch into five sub-routes. Each answer option triggers a more specific check that targets a potentially prohibited practice under Article 5:

- "Creates or expands a facial recognition database." → Q5 (Article 5(1)(e)).
  - In this case the AI system uses biometric data to create a facial database with the purpose of recognizing individuals.
- "Remote identification of individuals." → Q6, then Q7 (Article 5(1)(h) and Annex III(1)).
  - Identification is about discovering who someone is from a group of people. For example, if you have a photo of a face, you can try to find out who that person is by comparing the photo with a database of faces. 'Remote' means that this identification takes place without the active involvement of the person being identified. This option also covers tracking a natural person across cameras where the tracking relies on biometric data (e.g., gait, voice, face matched against a reference template), and body-worn cameras used for untargeted filming (e.g., of demonstrations).

> "For categorizing (labeling) people." → Q8 (Article 5(1)(g)).
  - This means that people are grouped or assigned a label based on certain characteristics, such as how they look or behave. For example, if an application labels people as 'shoplifters' or makes an estimate of someone's gender based on video images. Biometric categorisation does not include categorisation according to clothes or accessories (scarfs, hats, gloves, watches, crosses) or based on social media activity.
> "Recognizes or infers emotions or intentions." → Q9 (Article 5(1)(f)).
  - Emotions include happiness, joy, sadness, anger, aggression, surprise, disgust, shyness, excitement, shame, contempt, satisfaction, and pleasure.
  - **NOTE:** physical states such as pain or fatigue are not considered emotions (e.g., driver-drowsiness detection is not emotion recognition). The mere detection of expressions, gestures or movements (e.g., a smile) is not emotion recognition unless used to infer an emotion.
> "Other purpose." → directly to the Article 50 transparency scenarios (Q11).
  - AI systems that use biometric data for purposes other than identification, categorisation or emotion recognition (e.g., voice samples used for language or dialect analysis to substantiate a country of origin) may still be considered high-risk. The questionnaire will continue with follow-up questions to determine the correct risk category.

See also Recital 14 and 54 and considerations 120 and 121 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

Risk category **Q5 – Does this involve large-scale collection of facial images from the internet or security cameras (scraping)?**

Q5 is shown when Q4 is answered with 'Creates or expands a facial recognition database'. Article 5(1)(e) prohibits AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage. See also Recital 43 of the AI Act. If the answer is 'Yes', the questionnaire concludes that a prohibited practice is at stake and routes the user via the Article 2 exception screen to the Prohibited outcome. If the answer is 'No', the user is forwarded to the Article 50 transparency scenarios (Q11).

Risk category **Q6 – Is the AI system only used to confirm that someone is who they say they are (verification or authentication)?**

Q6 is shown if answer option 'Remote identification of individuals' is selected in Q4. If the answer is 'Yes', the AI system does not fall within the Article 5(1)(h) prohibition according to Recital 17 and the user is forwarded to the Article 50 transparency scenarios (Q11). If the answer is 'No', the user is forwarded to Q7 to assess the law-enforcement context. See also Recital 17 and 54 and considerations 129 through 144 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

To assist users, the following remark is provided:

Verification and authentication mean confirming an already-known identity: the person is present and willing and the system checks them against a reference that belongs to them. If the person is unknown, passive or not being matched against their own pre-stored reference, the system is not verifying or authenticating – it is trying to find out who someone is (identification) or infer something about them (categorisation). If the biometric data is not only used for authentication but simultaneously

compared to any other database (for example a criminal database used by Interpol or national police), this exception does not apply and 'No' should be answered here.

Examples of verification and authentication:
> Unlocking a phone by matching your face to the face stored on the device
> Fingerprint reader giving access to building by matching your fingerprint to your enrolled profile
> A camera checks if your face matches the photo in your passport.

Examples not involving verification or authentication, but identification or categorization:
> A camera at a stadium entrance scans every face in the crowd to find a wanted suspect (identification)
> An airport uses facial recognition to spot known terrorists among passengers without their knowledge (identification)
> A shop uses cameras to detect the emotions or stress levels of customers (categorisation)
> An HR system analyses job applicants' facial expressions during video interviews (categorisation)
> A system checks whether someone walking past a camera is male or female based on their face (biometric categorisation)
> A police database is searched with CCTV footage to identify who committed a crime (identification).

For verification or authentication to apply, the person must actively and consciously participate (e.g., step into a marked capture zone, present a fingerprint on demand). A surveillance camera in a metro station does therefore not fall under the verification or authentication, because the traveller is not actively participating in the identification. On the other hand, scanning a biometric metro ticket at a turnstile does, since the traveller actively participates in the verification process.

Risk category **Q7 – Is the AI system used for public safety, crime prevention, investigation or prosecution or the execution of criminal justice?**

Q7 is only shown when Q6 is answered with 'No'. It operationalises Article 5(1)(h), which prohibits real-time remote biometric identification in publicly accessible spaces for the purpose of law enforcement. See also Recital 32 and 33 and considerations 126, 127 and 128 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

If the answer is 'Yes' to Q7, the questionnaire routes the user to a question about exceptions for the prohibition of these systems (Q7-exceptions). If the answer is 'No', the user is forwarded to the cyber security and data protection exception (Q10).

To assist users, the following remark is provided:

Law enforcement entails among others investigation or prosecution or the execution of criminal justice. Exceptions for prohibition may apply, these will be discussed in the follow-up question Q7-exceptions.

Risk category **Q7-exceptions – Which of the following scenarios apply to your AI system?**

Q7-exceptions is only shown when Q7 is answered with 'Yes'. It operationalises Article 5(2) and Article 5(3) of the AI Act, which discuss conditions that must be met for the use of real-time biometric identification systems in publicly accessible spaces for the purpose of law enforcement to be authorised. See also Recital 34 and 35 and considerations 145 through 148 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

The following checkboxes are presented to the user:
> The AI system is used in the search for victims

of crime including missing persons or threats to the life or to the physical safety of natural persons or of a terrorist attack or the localisation or identification of perpetrators or suspects of specific criminal offences as listed in the user guidance below

> A fundamental rights impact assessment (FRIA) has been conducted for this AI system

> The AI system has been registered in the EU database

In duly justified cases of urgency, the use of such systems may be commenced without the registration in the EU database, provided that such registration is completed without undue delay.

> Authorisation for use has been granted by a judicial authority

A judicial authority should be understood in a broad sense. It can cover authorities that exercise judicial function or constitute part of the judicial system. Independent administrative authority whose decision is binding of the Member State in which the use is to take place is also considered a judicial authority. In a duly justified situation of urgency the use of the system may be commenced without an authorisation provided that such authorisation is requested without undue delay, at the latest within 24 hours.

If the user checks all the boxes and thus all conditions for exception are met, the questionnaire routes the user to the prohibited-with-exception outcome. If one of the conditions is not met, the exception does not apply and the questionnaire routes the user to the Article 2 exception screen on the prohibited path.

To assist users, the following remark is provided:

Please check the box if the AI system is used for the localisation or identification of perpetrators or suspects of the following criminal offences:

> terrorism
> trafficking in human beings
> sexual exploitation of children and child pornography
> illicit trafficking in narcotic drugs or psychotropic substances
> illicit trafficking in weapons, munitions or explosives
> murder, grievous bodily injury
> illicit trade in human organs or tissue
> illicit trafficking in nuclear or radioactive materials
> kidnapping, illegal restraint or hostage-taking
> crimes within the jurisdiction of the International Criminal Court
> unlawful seizure of aircraft or ships
> rape
> environmental crime
> organised or armed robbery
> sabotage
> participation in a criminal organisation involved in one or more of the offences listed above.

Risk category **Q8 – Does the AI system estimate race, ethnicity or origin, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation?**

Q8 is shown when Q4 is answered with 'For categorizing (labelling) people'. Article 5(1)(g) prohibits the placing on the market, putting into service or use of biometric categorisation systems that categorise natural persons individually based on their biometric data to deduce or infer race, ethnic origin, political opinions, trade-union membership, religious or philosophical beliefs, sex life or sexual orientation. See also considerations 165 through 170 of the draft guidelines on the classification of high-risk AI systems of the AI Act. These types of characteristics are referred to as sensitive personal characteristics according to Article 9 GDPR. The lawful labelling or filtering of biometric datasets by law enforcement is not covered by this prohibition.

If the answer is 'Yes', the questionnaire routes the user via the Article 2 exception screen to the Prohibited outcome. If the answer is 'No', the user is directed to the cyber security and data protection exception (Q10).

To assist users, the following remark is provided:

This does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in the area of law enforcement.

Risk category Q9 – Is the AI system used in the workplace or in education?

Q9 is shown when Q4 is answered with 'Recognizes or infers emotions or intentions'. Article 5(1)(f) prohibits AI systems used to infer emotions of a natural person in the workplace or in education institutions, except where the AI system is intended to be put in place or to be put into service for medical or safety reasons. See also Recital 44 and considerations 171 through 177 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

If the answer is 'Yes', the questionnaire routes the user via the Article 2 exception screen to the Prohibited outcome. If the answer is 'No', the user is sent to the cyber security and data protection exception (Q10).

To assist users, the following remark is provided:

The work environment also includes recruitment and selection and the home work environment. Education also includes admission to education and online education.

Risk category Q10 – Is the system used solely for the purpose of enabling cybersecurity and personal data protection measures?

Q10 is only shown within the biometrics branch when 'No' has been answered to the following prohibited practices:

- Q6: The AI system (processing biometrics data) is not only used to confirm that someone is who they say they are (verification or authentication).
- Q8: The AI system (processing biometrics data) does not estimate race, ethnicity or origin, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation.
- Q9: The AI system (processing biometrics data) is not used in the workplace or in education.

AI systems that are solely used for the purpose of enabling cybersecurity and personal data protection measures are not considered to be high-risk AI systems. See Recital 54. This type of AI systems refers to systems whose only job is to defend IT systems or safeguard personal data. For example, tools that detect malware, block intrusions, filter phishing, control access, anonymize data or spot data leaks.

If Q10 is answered with 'Yes', the Recital 54 exception applies and the user is forwarded to the Article 50 transparency Scenarios (Q11).

If Q10 is answered with 'No', the Recital 54 exception doesn't apply and users are forwarded to the Article 2 exception screen on the high-risk path.

To assist users, the following remark is provided:

This refers to AI systems whose only job is to defend IT systems or safeguard personal data. For example, tools that detect malware, block intrusions, filter phishing, control access, anonymize data or spot data leaks.

6.3 Transparency obligations and prohibited non-consensual fake nude imagery

Risk category **Q11 – Does one of the following scenarios apply to the AI system?**

Q11 is displayed in all cases where the user has not already been redirected to the exception screen prior to the determination that the AI system might be prohibited. Q11 determines whether generative or interactive AI is at stake. For this type of AI systems, additional transparency obligations apply (see section 7). The user selects every applicable scenario from a multi-select list (e.g., direct interaction with natural persons, synthetic-content generation, emotion or biometric-categorisation use, content generation or none of the above). The four scenarios refer to Article 50(1)-50(4) of the AI Act.

If a scenario related to (synthetic) image generation is selected, the user is redirected to Q12 on non-consensual fake nude imagery. In all other cases, the user is forwarded to the Annex III domain gate (Q13). Selected scenarios add a transparency-obligation badge to the results page. Selecting answer option 'None of the above' adds no such badge.

To assist users, the following remark is provided:

As clarified in Recital 134 and Article 50(6), the same AI system can be subject to Article 50 transparency obligations and be classified as high-risk under Article 6.

Risk category **Q12 – Is the AI system used for non-consensual fake nude imagery?**

Q12 is only shown when answer option 'It generates synthetic audio, image, video or text content.' or 'It generates or manipulates image, audio, video or public-interest text content.' is selected in Q11. In Q12, it is determined whether image generation is used for removing clothing from images of people without consent, a process also referred to as 'nudification' – which is prohibited under the AI Act as a result of AI Omnibus amendments.

If answered with 'Yes', the user is directly redirected to terminal node 'Prohibited'. If 'No', the user continues to Q13.

To assist users, the following remark is provided:

Also referred to as 'nudification'. The process of removing clothing from images of people.

6.4 Annex III high-risk domains

Risk category **Q13 – Does the AI system fall within any of the following domains?**

Q13 is the entry point for the high-risk classification under Annex III(2) to Annex III(8). The user selects one of the below nine high-risk domains or 'None of the above'. An answer option should also be selected when the AI system plays a small or supporting role within a domain. Each domain triggers one or more domain-specific questions (described in the below sections). If none of the domains apply, the user is forwarded to horizontal Article 5 prohibitions (Q28). See also consideration 118 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

> Critical physical and digital infrastructure → Q14 and Q15.

Critical digital infrastructure includes internet exchange points, DNS service providers, top-level domain name registries, cloud computing services, data centre services, content delivery networks, trust service providers, electronic communication services, public electronic communication networks. Additionally, critical physical infrastructure includes road traffic, supply of water, supply of gas, supply of heating and supply of electricity. Nuclear-specific safety components fall outside the supply

of electricity domain and are thus not included here. 'Supply of gas' excludes the oil sub-sector (operators of oil transmission pipelines, oil production / refining / treatment facilities, storage and transmission, and central stockholding entities). 'Supply of electricity' excludes the hydrogen sector (operators of hydrogen production, storage and transmission).

- Education or vocational training → Q16.
  - E.g., placement and admission, evaluation and testing, but also control such as plagiarism detection and anti-cheating software.
- Employment or HR → Q17 and Q18.
  - Examples how AI systems are used in the context of employment are:
    - Recording hours, measuring KPIs and performance
    - Task distribution
    - Recruitment and selection
    - Inflow and progression (employment conditions, contractual terms)
    - Outflow (termination of contract)

  Also choose 'yes' when the application broadly falls under the above, even if the application plays a small role in an otherwise manual process.
- Government benefits or public services → Q19.
  - E.g., processes around benefits or allowances, such as application, administration, execution, enforcement, recovery, and objection. This also includes waiving fines, repayments, or granting a payment arrangement. Select this option if the application plays a small role in the process.
- Assessment of creditworthiness, life or health insurance, or emergency services → Q20.
  - Applies for the following sub-uses:
    - Creditworthiness assessment or credit scoring of natural persons, where used to determine access to essential financial services (e.g., bank accounts, payment services with credit, loans, credit-line/ credit-card extensions, mortgages, publicly-backed loans). Including payment arrangements for repaying a public debt (e.g., for taxes, fines, or a benefit being reclaimed).
    - Risk assessment or pricing of natural persons in life or health insurance, including private long-term care insurance, personal pension products and credit life insurance. Motor, home and travel insurance are not covered.
    - Evaluating or classifying emergency calls, dispatching emergency services, or establishing priority in such dispatch (e.g., police, fire, ambulance, healthcare patient triage).
- Law enforcement → Q21 and Q22.
  - Law enforcement includes the prevention, investigation, and prosecution of criminal offences, the execution of sentences, and the protection of public safety. Think of police, special enforcement officers, and the execution of (community) sentences by probation services. Other agencies, such as juvenile justice organizations and specific (inspection) services, can also engage in law enforcement. Also select this answer for organizations and institutions that support law enforcement agencies, such as a forensics institute and for activities carried out on behalf of law enforcement agencies.
- Migration, asylum or border control → Q23, Q24 and Q25.
  - This domain covers AI systems intended to be used by or on behalf of competent public authorities in the migration, asylum and border-control fields, such as border-control authorities, consular and visa authorities, asylum determining authorities and Union

agencies. AI systems used exclusively for search-and-rescue fall outside of this domain.

> Administration of justice or dispute resolution → Q26.
  - Professional matters, in which a judge makes a ruling, fall under the judiciary. Alternative dispute resolution includes disciplinary law, ombudsmen, arbitration, and complaints committees such as the Advertising Code Committee, the Council for Journalism, and the National Complaints Committee for Education.
> Elections or democratic processes → Q27.
  - E.g., AI systems used to optimize the targeting and ad delivery of political advertising.
> None of the above → horizontal Article 5 prohibitions (Q28).

Risk category **Q14 – Does the AI system fulfil a safety function within the management or operation of this infrastructure?**

Q14 is only shown when answer option 'Critical physical and digital infrastructure' is selected in Q13. Q14 asks whether the AI system fulfils a safety function within the management or operation of critical infrastructure such as gas, water, heating or electricity supply, road traffic, or critical digital infrastructure, as mentioned in Annex III(2). See also Recital 55 and considerations 180 through 192 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

If the answer is 'No', the AI system is not classified as high-risk under Annex III(2) and the user proceeds to the horizontal Article 5 prohibitions (Q28). If the answer is 'Yes', the user is redirected to Q15.

To assist users, the following remark is provided:

A safety component performs a function that directly protects the physical integrity of the critical infrastructure and the health and safety of persons or property. A safety function falls into one of six categories:

> Monitor and detect: E.g., AI detecting abnormal behaviour in critical infrastructure operation.
> Detect maintenance or inspection needs: E.g., AI detecting whether safety-related parts are worn and need replacement before they fail.
> Prevent harm: E.g., AI preventing critical infrastructure from starting up if abnormal behaviour is detected.
> Control or limit harm: E.g., AI controlling a specific behaviour or function of critical infrastructure and adjusting it accordingly.
> Mitigate consequences: E.g., AI that triggers a safe stop when a dangerous condition occurs.
> Control or supervise another safety system: E.g., AI supervising via sensors another system that performs a safety function.

The following AI systems are not safety components in critical infrastructure:

> AI-enabled traffic-flow optimisation systems based on real-time data that do not directly trigger safety-relevant changes.
> AI used for predictive maintenance that only generates recommendations to humans without triggering safety actions.
> AI used solely for cybersecurity purposes.
> AI systems performing a 'mere conduit' role - passive conveyance of traffic or data without control or decision-making over its content or destination.
> AI systems designed solely to prevent non-physical harm (e.g., moral harm, rights infringement) or damages caused by third parties through intermediation services.
> AI systems whose protective function is already provided by a separate redundant or backup system that directly protects the physical integrity of the infrastructure - in such cases the AI is not what stops the harm from materialising and is not the safety component.

> AI systems that are merely supportive, informational, organisational or optimisation-oriented and do not directly protect physical integrity.

Risk category **Q15 – Is the AI system used solely for cybersecurity purposes (e.g., defending IT systems, detecting threats, preventing intrusions)?**

Q15 is only shown when Q14 is answered with 'Yes'. Q15 asks whether the AI system is solely used for cybersecurity purposes. Pursuant to Recital 55 and consideration 187 from the draft guidelines, cybersecurity systems supporting critical infrastructure are not classified as high-risk under Annex III(2). Therefore, when the answer is 'Yes' here the system is not classified as high-risk under Annex III(2) and the user proceeds to the horizontal Article 5 prohibitions (Q28). A 'No' forwards the user to Q15-CER.

To assist users, the following remark is provided:

Examples of AI systems used solely for cybersecurity:

> An AI honeypot that proactively identifies and neutralises cyber threats in real time.
> AI that engages with potential attackers to learn new attack patterns.
> AI that detects unauthorised access (e.g., alerting affected customers before a stolen password is misused).
> AI that detects suspicious email addresses and identifies stolen data.

Risk category **Q15-CER – Is the AI system intended to be used by an entity identified as critical under the Critical Entities Resilience Directive?**

Q15-CER is only shown when Q15 has been answered with 'No'. Q15-CER asks whether the AI system is intended to be used by a critical entity as defined by the Critical Entities Resilience (CER) Directive. Pursuant to considerations 188 through 192 of the draft guidelines on the classification of high-risk AI systems of the AI Act, an AI system should only be considered high-risk if it is used by an entity identified as a critical entity by a Member State under the CER Directive. Therefore, a 'Yes' here will lead to the Article 2 high-risk exceptions and a 'No' continues to the horizontal Article 5 prohibitions (Q28).

To assist users, the following remark is provided:

For the defined critical entities, please refer to the 'CER' Directive.

Risk category **Q16 – In which of the following education contexts is the AI system used?**

Q16 is only shown when answer option 'Education or vocational training' is selected in Q13. Q16 lists four high-risk uses of AI in education and vocational training, as mentioned in Annex III(3), Recital 56 and considerations 217 through 235 of the draft guidelines on the classification of high-risk AI systems of the AI Act. The below options should also be chosen when the AI system plays a part in the process, without being the only factor.

> Allocation or admission of students to education, training or schooling
>> E.g., Automated admission systems review student applications, transcripts and test scores, school-assignment systems used by municipalities or regional authorities based on location, sibling attendance and capacity, AI systems matching job seekers or employees to apprenticeships programs. Includes assignment to specific programs or courses within an institution. AI systems that are not included are: chatbots providing general admission information and matching platforms that help students select educational programs themselves (no decision making involved).

> Assessing the level or quality of education someone receives
  - E.g., AI-powered tools recommending educational levels, systems using placement tests and skill assessment for vocational training, AI classifiers used in special education that recommend a placement and level of support. AI systems that are not included are: AI that provides personalised education recommendations to students that help them choose their own path, AI used for trend analysis in level of education.
> Evaluating learning outcomes (assessment leading to a grade, qualification, certification or progression decision)
  - E.g., AI-enabled grading systems that count towards a final evaluation, AI-enabled personal learning assistants whose reports are used by the teacher in their grading. AI systems that are not included are: adaptive learning systems that adjust difficulty, tutoring systems with personalised feedback, learning analytics platforms (i.e. AI-based learning tools that do not lead to a grade of any form).
> Monitoring or detecting unauthorised behaviour of students during tests (e.g., proctoring or anti-cheating software)
  - E.g., Real-time behaviour analysis systems used to detect cheating during exams, AI exam monitoring using facial recognition during in-person exams. AI systems that are not included are: plagiarism checks performed on homework or essays after submission, non-academic student monitoring for instance in cafeterias or hallways.
> None of the above

Selecting any of these forwards the user to the Article 2 high-risk exceptions. Selecting 'None of the above' continues to the horizontal Article 5 prohibitions (Q28).

To assist users, the following remark is provided:

This question applies only to AI systems that are intended to be used in the context of, or within, an educational or vocational training institution at any level: formal or non-formal, public or private, including adult education and continuing education. For AI systems that learners use on their own for informal learning, where the system does not lead to a recognised credential, certification or qualification, 'None of the above' should be selected.

'Assessing the level or quality of education that someone receives or has access to' should also be selected if the application contributes to the assessment. There does not need to be (fully) automated assessment.

Examples of evaluating learning outcomes or guiding the learning process (personalized learning) are:

> Evaluation of (multiple choice) tests
> Grading assistance
> Assessment of learning pace or style
> Personalization of the curriculum or learning process.

Risk category **Q17 – In which of the following employment areas is the AI system used?**

Q17 is only shown when answer option 'Employment or HR' is selected in Q13. Q17 lists three high-risk uses in recruitment or selection of natural persons, as mentioned in Annex III(4)(a) and considerations 243 through 254 of the draft guidelines on the classification of high-risk AI systems of the AI Act. One of the below options should also be chosen when the AI system plays a part in the process. There does not need to be fully automated assessment or decisions.

> Placing targeted job advertisements
  - E.g., AI systems selecting which user sees specific job vacancies on platforms based on user characteristics that go beyond

occupational requirements. This also includes placing targeted job advertisements for both targeted recruitment and writing targeted job descriptions. AI systems that are not included are: advertisements highlighting workplace culture that is not tied to a specific vacancy, ads personalised on objective criteria that are genuine and determine occupational requirements (e.g., a job needing a forklift permit shown to certified forklift drivers), ads based exclusively on the content of the website where the ad is placed with no targeting of individual users using personal data.

- Analysing or filtering applications
  - E.g., AI systems producing suitability scores, ranking candidates, assessing the compatibility of a candidate's profile or experience with selection criteria or generating competence profiles that are used to create shortlists. Any analysis that affects access to later recruitment stages is included here. AI systems that are not included are: AI systems performing only descriptive analysis without applying evaluation techniques, AI that generates job descriptions from a list of tasks and qualifications previously defined by a human. However, AI that itself generates the required qualifications or that evaluates the candidates against its own generated job description, is included here.
- Evaluating candidates
  - E.g., AI systems that score candidates based on written or oral responses to an online assessment, predictive modelling of candidate suitability, qualitative judgements derived from data. The assessment does not need to determine the hiring outcome directly. It is enough if it plays a role in the decision-making process.

And three high-risk uses in employment, as mentioned in Annex III(4)(b) and considerations 255 through 278 of the draft guidelines on the classification of high-risk AI systems of the AI Act:

- Decisions affecting terms of work-related relationships
  - E.g., AI systems driving decisions on pay, working time arrangements, leave entitlements, denial of a leave request, significant changes to regular working hours or access to career development. Decisions here include omissions: not being granted a leave request is covered. AI systems that are not included are: day-to-day operational adjustments that do not materially alter the rights or obligations under the contract (e.g., allocation of office space, choice of break within an assigned shift).
- Decisions affecting promotion and termination of work-related contractual relationship
  - E.g., AI systems driving decisions on dismissal, redundancy, non-renewal of fixed-term contract, promotion or grade elevation, deactivation of a platform account and lengthy or frequent suspensions producing equivalent effects. Decisions here include omissions: not being promoted is included. AI systems that are not included are: AI that merely records and flags to a manager the forthcoming lapsing of a fixed-term contract before its endpoint because the termination follows directly from the contract.
- Task allocation among employees
  - E.g., AI systems driving task allocation based on individual behaviour or personal characteristics, allocation of more complex assignment to higher scoring workers or withholding delivery slots from workers with lower acceptance rates. AI systems that are not included are: systems driving task allocations based exclusively on objective, neutral and external criteria tied directly to

the job requirements, such as availability indicated by the worker, proximity to the work location, possession of a required professional accreditation. Such criteria cannot be a proxy for nationality or ethnicity.

See also Recital 57 of the AI Act. Selecting any of these forwards the user to the Article 2 high-risk exceptions. Selecting 'None of the above' continues to Q18.

To assist users, the following remark is provided:

'Employees' has a broad definition. It includes employees, freelancers, independent contractors and other self-employed persons, solo self-employed persons in a situation comparable to workers, and persons providing services through digital platforms. The workplace is also understood broadly: any physical or virtual space where natural persons engage in tasks assigned by their employer or by the organisation they are affiliated to. Determining the distribution of tasks among employees includes administrative processes where hours, productivity, speed, breaks, performance, or behaviour are processed.

Risk category **Q18 – Does the AI system monitor or evaluate the performance or behaviour of employees?**

Q18 is only shown when answer option 'None of the above' is selected in Q17. In contrast to the closed list of Q17, Q18 is broader formulated to capture tools, especially record-keeping HR systems, that don't fit the Q17 verbs (e.g., workforce-analytics, productivity-tracking, shift/scheduling tools, sentiment dashboards). Recital 57 follows the same split. See also considerations 255 through 278 of the draft guidelines on the classification of high-risk AI systems of the AI Act. If 'Yes', the AI system might be high-risk and users are brought to the Article 2 high-risk exceptions. If 'No', users are redirected to horizontal Article 5 prohibitions (Q28).

To assist users, the following remark is provided:

> We mean employees in the broadest sense, including employees but also freelancers and employees hired through a third party (temporary, secondment).
> Think of administrative processes where hours, productivity, speed, breaks, performance, feedback scores or behaviour are processed. E.g., AI rating workers on number of deliveries per hour or other quantitative productivity measures; qualitative indicators (feedback scores, performance assessments, appraisals, rankings) feeding into employee files; AI assessing or downgrading workers for low shift-acceptance rates; AI assessing punctuality, adherence to workplace rules, responsiveness, patterns of interaction with colleagues.
> Not included here is: AI deployed solely to meet external legal or regulatory obligations (e.g., transaction logging for financial-market abuse rules), provided the same system is not also used for employment monitoring; AI deployed purely and exclusively for medical and safety reasons; AI deployed purely and exclusively to protect property or company assets, where not also used for employment monitoring; AI that flags potential issues only to the worker (not to the employer or anyone else in the employer organisation) to support task performance; AI that conveys aggregate performance data only to the worker (e.g., comparing their results to company averages), without notification to managers; AI used purely to track delivery operations or detect workflow mistakes where the system informs only the employee in charge, who then adopts any needed correction.

Risk category **Q19 – Is the AI system used to assess whether someone is eligible for a service, provision or benefit or to decide whether a service is granted, denied, limited, withdrawn or reclaimed?**

Q19 is only shown when answer option 'Government benefits or public services' is selected in Q13. Q19 asks whether the AI system is used to assess eligibility for a public service, provision or benefit, or to decide whether such a service is granted, denied, limited, withdrawn or reclaimed, as listed in Annex III(5)(a). See also Recital 58 and considerations 282 through 293 of the draft guidelines on the classification of high-risk AI systems of the AI Act. A 'Yes' forwards the user to the Article 2 high-risk exceptions. A 'No' continues to the horizontal Article 5 prohibitions (Q28). The question should be interpreted broadly, as in Q33 and Q34 it will be checked whether conditions for the high-risk 'filter' (narrow procedural tasks, aimed to improve the results of previously completed human activity etc.) apply.

To assist users, the following remark is provided:

The following applications fall under the mentioned service, provisions or benefits:

> The benefits or services involved are essential — i.e., necessary for natural persons to fully participate in society or improve their standard of living. The illustrative categories are:
  - i) healthcare services including long-term care;
  - ii) social-security benefits (sickness, maternity / paternity, invalidity, old-age, survivors', accident-at-work and occupational-disease, death grants, unemployment, pre-retirement, family, long-term-care, minimum-income, housing benefits, benefits to access essential services such as water, energy, digital communications);
  - iii) social services (social assistance, employment and training services, housing support and social housing, childcare).

> The AI system evaluates eligibility for those benefits, or is used to grant, deny, reduce, revoke or reclaim them — including as a preparatory step (e.g., selection for control).

Not included here is e.g., AI assessing applications or claims from companies (not natural persons); AI that only matches persons with the support services that fit them (without evaluating eligibility); AI for case-handler allocation; AI for proactive identification of persons in need of preventive care (where outputs are not used to prioritise / exclude / influence access). Tax remission and analogous waivers of public revenue are also outside.

Risk category **Q20 – Is the AI system used for any of the following?**

Q20 is only shown when answer option 'Assessment of creditworthiness, life and health insurances or emergency services' is selected in Q13. Q20 lists three high-risk uses as listed in Annex III(5)(b)-(d) and Recital 58 of the AI Act. See also considerations 294 through 336 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

> Creditworthiness assessment or credit scoring of individuals
  - E.g., AI systems evaluating a person's creditworthiness or establishing a credit score, where used to determine access to essential services (such as a bank account, payment services where credit is provided as a feature, loans and credit, extensions of credit-line or credit card limit, mortgages, publicly-backed financial services). AI systems that are not included are: AI intended only for pricing of credit without itself evaluating creditworthiness or a credit score, AI used for non-essential services such as stock, trading, complex financial instruments, premium credit cards or travel

loans, AI mainly used for fraud detection, anti-money-laundering or internal prudential purposes.

> Risk assessment or pricing for life or health insurance

E.g., AI systems used for risk assessment or pricing in relation to natural persons in life or health insurance such as private long-term care insurance, personal pension products (that can impact livelihood at older age), credit life insurance, health insurance offered by publicly-finance insurers, privately serviced health insurance even where the Member State has a universal public healthcare system. AI systems that are not included are: systems used for risk assessment or pricing of other insurance lines and insurance-based investment products, AI used for claims management and product design. AI systems used for non-essential services (e.g., motor or travel insurance, complex financial instruments, leisure loans) are also not included here. Please note: AI used for fraud detection is not excluded here. Only when the fraud detection feature is a distinct system that feature will be excluded, the risk-assessment or pricing system it is a part of is still included here.

> Evaluating emergency calls, deploying or prioritising emergency services or triage of urgent medical care

E.g., AI systems used for one of the following: evaluating or classifying emergency calls by natural persons, AI used to dispatch emergency first response services, AI establishing priority in such dispatch, AI dispatching emergency healthcare patient triage systems, AI establishing priority in triage dispatch. AI systems making forecasts in such situations is only included if it is specifically intended to decide when or where to dispatch emergency first-response services. Generic forecasting without a dispatch link is not included. Additionally, AI systems that are not included here are: AI transcribing emergency calls, general disaster forecasting, AI training simulator for first responders, AI for patient identification at medical procedures, AI for call-traffic management of the emergency communication system, AI used to estimate healing time, bed management or appointment scheduling.

> None of the above

Selecting any of these forwards the user to the Article 2 high-risk exceptions. Selecting 'None of the above' continues to the horizontal Article 5 prohibitions (Q28).

To assist users, the following remark is provided:

Emergency services include police, fire brigade, and ambulance. Credit scoring also includes determining if someone can get a payment arrangement for taxes, fines or repaying a benefit.

Choose 'None of the above' if the AI system is used for detection of financial fraud or for non-essential financial services.

Risk category **Q21 – Is the AI system used in assessing the risk that someone will (re)commit a criminal offense?**

Q21 is only shown when answer option 'Law enforcement' is selected in Q13. Q21 asks whether the AI system is used to assess the risk that someone will (re)commit a criminal offence. See also Recital 42, Recital 59 and considerations 266 through 268 of the draft guidelines on the classification of high-risk AI systems of the AI Act. A 'Yes' routes the user to Q21 follow-up. A 'No' continues to Q22.

To assist users, the following remark is provided:

Also think of offenses and fraud risk scores, where an estimate is made of whether someone might commit a criminal offense and an estimate of the likelihood of recidivism. AI systems specifically intended to be used for administrative activities by tax and custom authorities as well as financial intelligence units carrying out administrative tasks analysing information pursuant to anti-money laundering law, are excluded here. Fraud-risk scoring AI is covered only if the underlying conduct constitutes a criminal offence; pure administrative fraud-risk scoring is not.

Risk category **Q21 follow-up – Is the risk assessment based solely on profiling of the natural person or on assessing personality traits and characteristics?**

Q21 follow-up is only shown when Q21 is answered with 'Yes'. Article 5(1)(d) only prohibits AI systems that assess the risk that someone will (re)commit a criminal offense when this is solely based on profiling or personality assessment. Otherwise, the system is considered to be high-risk under Annex III 6(d). This is echoed in considerations 366 through 368 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

If Q21 follow-up is answered with 'Yes', the user is redirected to Article 2 exception screen on the prohibition path. If 'No', the user is forwarded to Article 2 high-risk exceptions.

To assist users, the following remark is provided:

AI systems assessing (re)offending of criminal offences that are solely based on profiling or assessing personality traits are prohibited. Choose 'Yes' if the AI system assesses the risk of offending or re-offending using profiling, personality traits, past criminal behaviour, or any combination of these as a basis.

If the assessment also relies on objective and verifiable facts directly linked to a criminal activity (e.g., prior convictions, behaviour during custody, evidence in an open investigation), choose 'No'.

Risk category **Q22 – Does the AI system do any of the following?**

Q22 is only shown when answer Q21 is answered with 'No'. Q22 lists five high-risk uses in law enforcement, as stated in Annex III(6). See also considerations 356 through 370 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

> Assessing or estimating that someone will become a victim of crime
  - E.g., domestic violence risk assessment systems analysing victim statements, police reports, previous incidents and restraining order data, human trafficking vulnerability detection systems. AI systems that are not included are: AI systems predicting location-focused risks, road traffic accident prediction, natural catastrophe predictions.
> Acting as a polygraph or similar instrument
  - E.g., AI enabled lie detectors and similar tools such as voice stress analysis, eye tracking, brain imaging deception detection, behavioural analysis of micro expressions, body language or speech cues for credibility assessment during interrogation. AI systems that are not included are: AI monitoring police officer's fatigue or stress, plain interview transcription with tone flagging that does not draw truthfulness conclusions, AI that flags irregularities across polygraph examiners' interpretations.
> Assessing the reliability of evidence
  - This includes AI systems whose core function it is to evaluate the evidentiary reliability. This may include authenticity verification, data integrity checks, source reliability assessment and evaluation whether the evidence aligns with other available information

or is internally consistent. Examples are: AI-enabled digital forensics solutions for image, audio and video analysis, AI systems intended to be used to determine and evaluate alterations of documents or forged signatures, AI systems used for evaluating the authenticity of digital evidence retrieved from mobile or other devices including by analysing digital traces. AI systems that are not included are: systems used for working with the information collected but not used for the evaluation of the reliability of the evidence such as image enhancement and data recovery, reconstructing crime scenes, pattern detection, content flagging and searches.

> Automatically evaluating or predicting characteristics, behaviour or personal aspects (profiling) in the context of crime and assessing personality traits, characteristics or previous criminal behaviour

  E.g., AI systems assisting the identification of potential suspects, AI systems used for targeted online monitoring. AI systems that are not included are: systems analysing neighbourhoods and assigning criminality score where they do not include profiling of individuals, systems detecting suspicious transaction patterns that indicate money laundering or terrorist financing, even though that allows linking such transaction to personal profiles, AI-enabled crime-linkage systems analysing police crime reports and case files with the purpose of identifying patterns and similarities across different incidents.

> None of the above

See also Recital 59 of the AI Act. Any of these forwards the user to Article 2 high-risk exceptions. 'None of the above' continues to horizontal Article 5 prohibitions (Q28).

To assist users, the following remark is provided:

The concept of 'law enforcement authority' here includes police, public prosecutors, probation and parole services, prison services and customs authorities when acting in their criminal investigation capacity. When the AI system is not applied directly to natural person or does not have the capacity to materially influence decision-making choose 'None of the above'. Examples of such systems are systems used for automating administrative tasks, chatbots on police websites, automated number plate readers seeking to identify vehicles, systems checking whether road safety rules are broken etc.

AI systems that are not included here are: AI used in administrative proceedings (tax administration, customs in administrative capacity for import / export / transit procedures, financial-intelligence units performing administrative tasks under anti-money-laundering law). AI used by national intelligence authorities outside law-enforcement activities (excluded from the AI Act under Article 2(3)); AI systems often used by law enforcement that are not directly applied to natural persons or have the capacity to materially influence decision-making; Administrative AI (case management, accounting, equipment management, shift scheduling); Chatbots on police websites providing general information; Automated number-plate readers (vehicle-focused); Road-safety rule detection (speeding, helmet use, phone use while driving); Object detection (weapons, suspicious parcels, stolen goods, cultural objects); Detection of unusual behaviour or movements without identifying persons (running, raised arms, fighting stance); Gunshot-sound detection; Customs AI for goods assessment based on goods / routing / payment data.

Choose 'None of the above' if your system fits one of these out-of-scope categories.

Risk category **Q23 – Does the AI system do any of the following?**

Q23 is only shown when answer option 'Migration, asylum or border control' is selected in Q13. Q23 lists three high-risk uses, as stated in Annex III(7). See also Recital 60 and considerations 388 through 400 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

> Acting as a polygraph or similar instrument
  - AI enabled lie detectors and similar tools such as voice stress analysis, eye tracking, brain imaging deception detection, behavioural analysis of micro expressions, body language or speech cues for credibility assessment used to indicate the credibility of the person. AI systems that are not included are: systems that process interview audio but do not infer deception, systems only intended to create neutral, non-prioritised summaries of interviews, without credibility scoring or recommendations.

> Assessing risks for a person entering EU territory
  - E.g., person-level migration or security risk scoring, person-level health-risk flagging, object and itinerary signals mapped to person-level risk, group-level risk indicators applied to individuals, AI system consisting of several non-high-risk components that taken together assess the risk posed by a natural person. AI systems that are not included are: aggregate or cohort-level analytics, vehicle-focused screening without mapping to a person, post-decision data cleaning.

> Processing asylum, visa or residence-permit applications or related complaints
  - E.g., evidence appraisal and consistency analysis, automated comparison of discrepancies against prior submissions, origin inference used in the examination. AI systems that are not included are: systems that provide applicants with answers to frequently asked questions, to guide users to the correct forms, to schedule appointments or to deliver application status notifications. Such an AI system does not assist the authority in examining eligibility or associated complaints, but its intended purpose is service delivery to migrant applicants.

> None of the above

Any of these forwards the user to the Article 2 high-risk exceptions. 'None of the above' continues to Q24.

Processing asylum, visa or residence permit applications or handle complaints related to the eligibility of the applicant is also applicable when the AI system is used to detect, recognise or identify individuals, or for assessing the reliability of evidence for applications or complaints.

To assist users, the following remark is provided:

These AI systems should be intended to be used by or on behalf of all national authorities designated to pursue the EU or national law regarding the field of migration, asylum, and border-control management, as well as Unition institutions, bodies, offices or agencies acting within their mandates under that law.
AI systems used exclusively for search-and-rescue (lifesaving) purposes are excluded and 'None of the above' should be selected.

Risk category **Q24 – Is the AI system used for the detection, recognition or identification of individuals?**

Q24 is only shown when answer option 'None of the above' is selected in Q23. Q24 asks whether the AI system is used for detection, recognition or identification of individuals in a migration or border-control context. A 'No' exits the high-risk

path and the user is redirected to horizontal Article 5 prohibitions (Q28). See Annex III(7), Recital 60 and considerations 401 through 403 of the draft guidelines on the classification of high-risk AI systems of the AI Act. If 'Yes' is selected, whether a high-risk classification of the AI systems is applicable is determined in Q25.

To assist users, the following remark is provided:

Examples are: identification of persons at border-crossing points, detection that cues operational action at land or sea borders, person-level monitoring in controlled facilities. AI systems that are not included are: AI used exclusively for search-and-rescue (lifesaving), with outputs not cueing border-control units. Dual-use systems (border control + SAR) are in scope by virtue of the border-control intended use; AI for travel-document / identity-document verification; AI in maritime or unmanned platforms limited to collision avoidance / safety of navigation (no border-control person cues); Analytics on historical data identifying document-fraud patterns at aggregate level (unless the patterns are then applied to named individuals, then in scope); Crowd-size or gate-occupancy estimation that does not detect, recognise or identify persons and does not generate person-level alerts.

Risk category **Q25 – Is the AI system solely used to verify a passport or other travel document to confirm that a specific person is who they claim to be?**

Q25 is only shown when Q24 is answered with 'Yes'. Q25 asks whether the AI system is used solely to verify a passport or other travel document. A 'Yes' here brings users to horizontal Article 5 prohibitions (Q28). A 'No' forwards the user to the Article 2 high-risk exceptions. See Annex III(7), Recital 60 and consideration 380 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

To assist users, the following remark is provided:

If the AI system is used solely to verify a travel or identity document, including checking for morphing, unusual inks, copied signatures, chip or Public Key Infrastructure (PKI) validity, choose 'Yes'.
Choose 'No' when the AI system is also used for identification that is broader than just checking a passport or other travel document. Also choose 'No' whenever the AI system performs identification, detection or tracking that goes beyond confirming the holder of a presented document is the document's owner.

Presentation-attack checks / on-person liveness checks need careful consideration:

> If it is used solely as an anti-spoofing safeguard within a one-to-one verification flow at a seamless border procedure (e.g., an e-gate where the traveller actively presents themselves, choose 'Yes'
> If it is used to support operational person detection, tracking or identification beyond verification, choose 'No'.

Risk category **Q26 – Is the AI system used only for supporting administrative activities that do not affect the judiciary in individual cases?**

Q26 is only shown when answer option 'Administration of justice or dispute resolution' is selected in Q13. Q26 asks whether the AI system is used only for supporting administrative activities that do not affect the judiciary in individual cases (such as anonymising decisions or supporting communication between staff), as stated in Annex III(8)(a). See also Recital 61 and considerations 405 through 434 of the draft guidelines on the classification of high-risk AI systems of the AI Act. A 'Yes' brings users to horizontal Article 5 prohibitions (Q28). A 'No' forwards the user to the Article 2 high-risk exceptions.

To assist users, the following remark is provided:

Think of anonymizing or pseudonymizing judicial decisions, documents or data, communication between staff members, and administrative tasks. If the AI system is used to assist in researching and interpreting facts and the law and in applying the law to a concrete set of facts, the application is broader than just administrative activities and the answer should be 'No'.

Risk category **Q27 – Is the AI system intended to be used to influence the outcome of an election or referendum, or to influence the voting behaviour of natural persons (including their decision to vote, abstain, or cast a blank or invalid ballot) in the exercise of their vote?**

Q27 is only shown when answer option 'Elections or democratic processes' is selected in Q13. Q27 asks whether the AI system is used in elections or referenda in a way that can influence the outcome or the voting behaviour of natural persons, as stated in Annex III(8)(b). A 'Yes' forwards the user to the Article 2 high-risk exceptions. A 'No' brings users to horizontal Article 5 prohibitions (Q28). See also Recital 62 and considerations 435 through 447 of the draft guidelines on the classification of high-risk AI systems of the AI Act.

To assist users, the following remark is provided:

This does not include AI systems to the output of which natural persons are not directly exposed, such as tools used to organise, optimise or structure political campaigns from an administrative or logistical point of view.

Examples include: AI-enabled targeting of political ads; AI-enabled chatbot to promote support for candidates or policies; AI-enabled voter advice applications.

AI systems that are not included here are: AI used to organise / optimise / structure political campaigns from an administrative or logistical point of view: campaign-staff scheduling, rally logistics, topic selection, donor-database analysis; AI generating political-ad content (e.g., drafting slogans) without further dissemination; AI monitoring elected officials' legislative submissions for informational purposes; AI for sociological polling / research that informs the public without seeking to influence electoral processes; AI for academic electoral-behaviour modelling; AI for ballot counting; AI election-information chatbots operating in a politically neutral way for election authorities; General-purpose AI systems and content recommender systems whose recommendations may incidentally include political content but are not intended to influence elections, provided sufficient safeguards (e.g., refusing voting advice when asked).

6.5 Article 5 horizontal prohibitions

Risk category **Q28 – Does the AI system use techniques people cannot consciously perceive (e.g., subliminal techniques), or is it designed — or known to unintentionally — manipulate people's behaviour or choices without their awareness?**

Q28 is shown if an AI system has exited the high-risk path but could still potentially qualify as a prohibited application under Article 5. Under Article 5(1)(a), an AI system is prohibited if it uses techniques people cannot consciously perceive – such as briefly displayed subliminal images – or is designed (or known) to manipulate people's behaviour or choices without their awareness, whether intentionally or not. If the answer to Q28 is 'No', the user is forwarded to Q30. If the answer is 'Yes', Q29 asks whether the AI system can potentially have negative consequences for the user or others, which is a further requirement under Article 5(1)(a) for the system to qualify as prohibited.

To assist users, the following remark is provided:

Subliminal techniques include showing images so briefly that they cannot be consciously perceived. Manipulation without awareness includes AI chatbots that give advice influencing decisions, applications that stimulate addiction or systems that prevent users from filing a complaint while they intended to.

Risk category **Q29 – Can the AI system potentially have negative consequences for the user or other people?**

Q29 is shown if an AI system Q28 is answered 'Yes'. Q29 is the second-stage test whether the potential prohibition of the AI system under Article 5(1)(a) is applicable, i.e., distort behaviour in a manner that "causes or is reasonably likely to cause significant harm". A 'Yes' to Q29 routes the user to Article 2 exceptions on the prohibited path. A 'No' brings the user to Q30.

To assist users, the following remark is provided:

Think of an AI health chatbot intended to give tips on a healthy lifestyle and tailored advice for psychological and physical exercises. However, if the chatbot encourages unhealthy habits or dangerous activities (e.g., excessive exercise without rest or drinking water) that they would not have done otherwise, it has negative consequences for the user.

Risk category **Q30 – Does the AI system evaluate, score or classify individuals based on their social behaviour or personal characteristics in a way that could lead to unfavourable treatment?**

Q30 is shown when Q28 or Q29 are answered with 'No', i.e., no prohibited subliminal or manipulative techniques are applicable. Q30 asks whether the AI system might still be prohibited under Article 5(1)(c) – also referred to as 'social scoring'. Whether social scoring is applicable is also a two-step exercise. If Q30 is answered with 'No', the questionnaire concludes that 'No requirements' follow from the AI Act. If Q30 is answered with 'Yes', users are directed to a follow up question (Q31) to determine whether prohibited social scoring is applicable.

To assist users, the following remark is provided:

Examples of social behaviour are:

> Driving behaviour
> Communication style
> Breaking (un)written rules

Examples of personal characteristics are:

> Age
> Height
> Gender
> Economic situation
> Movement patterns
> Interests

Risk category **Q31 – Does the unfavourable treatment occur in a completely different context from the behaviour or characteristic it is based on or is it disproportionate to that behaviour?**

Q31 is shown if Q30 is answered with 'Yes'. Q31 checks whether the AI system leads to detrimental or unfavourable treatment "in a context unrelated to the data origin" or "unjustified or disproportionate to the gravity", which is a condition of the prohibition of Article 5(1)(c). A 'Yes' to Q31 routes users to Article 2 exceptions on the prohibited path. A 'No' concludes that 'No requirements' follow from the AI Act.

To assist users, the following remark is provided:

> Unfavourable treatment in an unrelated context is, for example, a fraud risk score that leads to an

investigation, where the fraud risk score is based on characteristics without clear relevance to the assessment of fraud, such as having a partner with a certain nationality or behaviour on social platforms.

> Disproportionate treatment is, for example, someone getting a higher fraud risk score because municipal taxes were once paid late or a student not being admitted to further education because they were once caught cheating.

6.6 Article 6(3) – Profiling and limited tasks

Risk category **Q32 – Is the AI system used to automatically evaluate or predict characteristics, behaviour or personal aspects?**

Q32 is only shown when AI systems are potentially high-risk and no exceptions apply ('None of the above' is selected for the Article 2 exception question). Q32 asks whether the AI system is used to automatically evaluate or predict characteristics, behaviour or personal aspects of natural persons (i.e., profiling within the meaning of Article 4(4) GDPR). Article 6(3) provides that an AI system referred to in Annex III is not high-risk if it does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons. However, this carve-out is not available when the AI system performs profiling. See also Recital 53 and considerations 109 through 112 of the draft guidelines on the classification of high-risk AI systems of the AI Act. A 'Yes' to Q32 therefore yields the high-risk outcome. A 'No', users are redirected to Q33.

If profiling is determined to be applicable in the Identification questionnaire and the user answers "No" to Q32, a warning is displayed: "According to questionnaire Identification, we expect this answer to be 'Yes'. Please reconsider carefully whether the AI system is used to automatically evaluate or predict characteristics, behaviour or personal aspects."

To assist users, the following remark is provided:

Think of automatically evaluating professional performance, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements of a person to analyse or predict. The evaluation must be carried out based on personal data. Examples of AI systems that do not perform profiling are:

> Classification or prediction based purely on non-personal data (e.g., goods, weather, vehicle telematics not tied to a natural person). Example: a customs AI assessing the risk of goods entering the EU on the basis of container number, description, routing, transport and payment method. It evaluates compliance risk of goods, not personal aspects of a natural person.

> Simple classification of individuals on personal characteristics (age, sex, height) without prediction, inference or drawing of conclusions.

NOTE: the profiled subject can be on either side of a decision. E.g., an AI that analyses recruiters' past decision-making patterns and evaluates their personal characteristics is profiling (the recruiters are the subjects), even though the AI's headline purpose concerns the candidates.

Risk category **Q33 – Does the AI system fall within one of the following scenarios?**

Q33 is only shown when Q32 is answered with 'No' (profiling is not applicable). Q33 asks whether the AI system is used only for narrow procedural or preparatory tasks:

> A narrow procedural task

E.g., AI systems used to categorise, change the format, structure or presentation of data or change its metadata. However, when an AI system performs a value judgement of data relevant for decision-making, for example categorisation of input data as 'useful' or 'less useful' for the human assessment, or attributing a score or ranking to input data,

are likely to have a broader impact on the assessment process and will therefore not be considered to perform only a 'narrow procedural task'.

> The improvement of the result of a previously human-performed activity
  - E.g., systems that flag errors or contradictions in finalized human ork that perform a quality-assurance function, systems that map conclusions to evidentiary records to strengthen the traceability of a decision without substituting the human judgment, systems that convert human-validated content for interoperability or accessibility purposes that transform but do not replace human output.
> The detection of decision-making patterns or deviations from previous decision-making patterns
  - E.g., an AI system that, after a set of decisions has been made, analyses patterns in decision-making so that human decision-makers can learn from them.
> A preparatory step that does not influence the outcome
  - E.g., an AI system that assesses data relevant to a decision and provides a human operator with references to the relevant legal provisions, information on jurisdiction and possibly existing internal guidelines relevant to the decision-making process.
> None of the above

When answer option 'Yes, the AI system is used to detect decision-making patterns or deviations from previous decision-making patterns' is selected, the user is brought to Q34 to check if conditions for this exemption apply. See also considerations 97 through 102 of the draft guidelines on the classification of high-risk AI systems of the AI Act. If one of the other answer options other than 'None of the above' are selected, the user is redirected to the high-risk-with-exception outcome (Article 6(3)). Selecting 'No' yields the high-risk outcome.

According to considerations 125 through 116 of the draft guidelines on the classification of high-risk AI systems, an assessment should be performed prior to applying the filter mechanism. This assessment should contain a description of the intended purpose of the AI system, a description why the systems qualifies as high-risk pursuant to Article 6(2) AI Act, a description which condition or conditions listed under Article 6(3) AI Act are considered to apply and why and a description why the system does not perform profiling, in which case the filter mechanism of Article 6(3) AI Act should not be applied. These criteria have already been assessed in preceding stages of the AI AQT questionnaires and therefore the assessment has indirectly already been performed. In particular:

> The intended purpose is addressed in metadata requested about the application in the tool
> The qualification of high-risk under Article 6(2) is addressed in the Risk classification questionnaire
> The conditions listed under Article 6(3) considered to apply are addressed in **Q33**
> The description why the system does not perform profiling is addressed in Q5 of the Identification questionnaire.

Therefore, the answers from these preceding questions can be used as the assessment for the filter mechanism and no additional explanation will be required from the user.

Moreover, when the exception is based on the AI system being part of a preparatory step, according to considerations 103 to 108 of the draft guidelines on the classification of high-risk AI systems of the AI Act, it must be true that:

> The system's output is general input or supplementary information for the human assessor, not a specific recommendation or evaluation of the case
> The possible impact of the system's output is very low in terms of representing a risk for the assessment to follow.

However, these conditions have already been touched upon Q8 through Q12 of the Identification questionnaire and are thus not reiterated to the user here.

To assist users, the following remark is provided:

The mere fact that a human is involved in procedural tasks isn't enough to fit any of the scenarios listed above. If the AI system is part of a more complex or agentic configuration where multiple AI components together materially influence individual decisions in a high-risk use case, this exception cannot be applied to individual modules in isolation. More information about meaningful human intervention in algorithmic-driven decision-making procedures can be found in another white paper of Algorithm Audit.[35]

Risk category **Q34 – Which of the following scenarios apply?**

Q34 is only shown when answer option 'Yes, the AI system is used to detect decision-making patterns or deviations from previous decision-making patterns.' is selected in Q33. Q34 asks whether the following scenarios apply:

> The human assessment is completed
  - This means that all necessary steps in the evaluation process have been taken.
> The AI system only performs an ex-post comparative assessment, which involves determining whether a current decision aligns with the previous patterns or decisions
  - The AI system must not infer relevant criteria from previous decisions and propose a new assessment based on those criteria.
> The AI system does not replace or influence the previously completed human assessment without proper human review
  - 'Proper' means that the human review applied after the AI system's assessment should be meaningful and comprehensive and it should consider, together with the result provided by the AI system, the various elements considered during the previously completed human assessment. Therefore, the output of the system can highlight discrepancies with previously made decisions but cannot modify the input or alter the underlying reasoning. To allow proper human review, the system should not only flag but also provide information on the discrepancies. It should however merely inform (e.g., through analytical tools used for quality assurance and statistics) and not replace the human assessment.

The three scenarios displayed resemble the conditions that should be met for this exception to apply. See considerations 97 through 102 of the draft guidelines on the classification of high-risk AI systems of the AI Act. If all scenarios are selected, the user is directed to the high-risk-with-exception outcome (Article 6(3)). If any one of the conditions is not met, the exception does not apply and the user is directed to the high-risk outcome.

To assist users, the following remark is provided:

Example: An AI system intended to analyse past eligibility checks completed by public administrators in the public administration in order to detect decision-making patterns or deviations for quality-assurance and reporting, without proposing outcomes on live cases or evaluating the performance of staff members, e.g., in the context of the annual appraisal.

[35] Public standard 'Meaningful human intervention for risk profiling algorithms', Algorithm Audit (2025).

6.7 Article 2 exception

Risk category **Article 2 exception – Does one of the following exceptions apply to your application?**

The Article 2 exception question asks whether the AI system falls under one of the Article 2 exceptions:

> Scientific research and development (Article 2(6))
  - This includes AI systems developed by students as part of their studies or by scientists affiliated with a university. In that case, the research must follow recognized ethical and professional standards for scientific research. For example, all participants must be informed that they are part of scientific research.
> Defence or military purposes (Article 2(3))
  - The type of entity carrying out the activities here does not matter. As long as the AI system is used in activities carried out for defence or military purposes.
> National-security purposes (Article 2(3))
  - National security purposes mean preventing and punishing activities that can seriously destabilize the constitutional, political, economic or social structures of a country. Think of direct threats to society, the population or the state itself, such as terrorist activities.
  - **NOTE:** When the AI system is also used for other activities or purposes, such as regular prevention, detection, and investigation of criminal offenses, choose 'None of the above'.
> None of the above

When the AI system is used for operational processes or regular business operations, the user should select 'None of the above'.

If a research exception is selected, the user is asked to describe the research (purpose, department, partners). The outcome in all three exception cases is the corresponding high-risk-with-exception or prohibited-with-exception, depending on the path that led to the Article 2 exception question. If 'None of the above' is selected and the user is on the high-risk path, the questionnaire continues to Q32 (Article 6(3) profiling check). If the user is on the prohibited path, 'None of the above' yields the Prohibited outcome.

Flowchart Risk classification

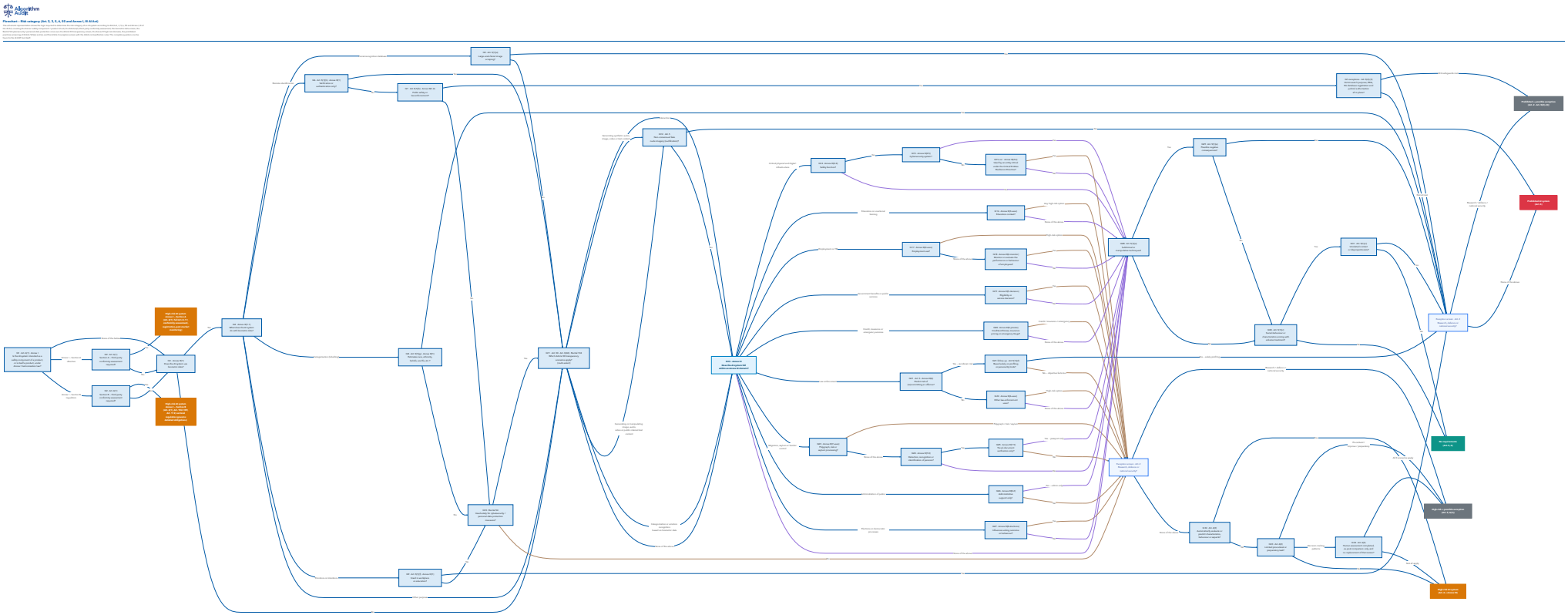


Figure 28 - Flowchart determining the risk category of the AI system.

7. Questionnaire Obligations

The questionnaire Obligations describes the requirements that follow from the AI Act depending on the role, status and risk category of the AI system. The questionnaire consists of three questions. When the questionnaires Role and status and Risk category have been completed, the answers are pre-filled automatically; they can also be selected manually.

- **Q1 – What role do you have in relation to the AI system?** The role determines which set of obligations applies. The AI Act distinguishes provider, deployer, importer, distributor and authorised representative. A private user, using an AI system in the course of a purely personal, non-professional activity, has no obligations under the AI Act. The roles provider and deployer can be held simultaneously; the other roles are treated as mutually exclusive.
- **Q2 – What is the usage status of the AI system?** The status (in use or in development) determines the compliance deadline that applies.
- **Q3 – What is the risk category of the AI system?** The risk category determines the substantive obligations. The categories are no requirements, prohibited, high-risk, generative and interactive AI, and exception. The categories high-risk and generative and interactive AI can apply at the same time.

The obligations follow from Chapter III of the AI Act (Article 6, read in combination with Annex I or Annex III) and from Article 50 of the AI Act. For AI systems classified as 'no requirements', no obligations follow from the AI Act. Where an exception applies, for example for scientific research and development as referred to in Article 2, the obligations are those of the underlying category to which the exception attaches, e.g., law enforcement exceptions for prohibited usage of biometric identification in public spaces moves down the system to the high-risk regime rather than out of the regulation entirely. The sections below describe the obligations per risk category: prohibited AI systems (section 7.1), high-risk AI systems (section 7.2) and generative and interactive AI (section 7.3).

7.1 Obligations for prohibited AI systems

An AI system that meets the definition of a prohibited AI practice (Article 5) may not be placed on the market, put into service or used within the European Union. This prohibition has applied since 2 February 2025. As of 2 December 2026, a new category of prohibited AI systems as introduced by the AI Digital Omnibus becomes applicable. This entails AI systems generating non consensual intimate images, videos, audio and similar material and child sexual abuse material. There is no compliance pathway for prohibited AI systems: the system must be discontinued.

Provider

This AI system must be taken out of service immediately. Its intended use constitutes a prohibited AI practice under the EU AI Act and may not be placed on the market, put into service or used within the European Union.

Deployer

Stop using the AI system. Inform the provider that the system must be taken out of service immediately. Its intended use constitutes a prohibited AI practice under the EU AI Act and may not be placed on the market, put into service or used within the European Union.

Further guidance is available in the Guidelines on Prohibited Artificial Intelligence[36] published by the European Commission.

[36] Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689.

7.2 Obligations for high-risk AI systems

High-risk AI systems must comply with the requirements set out in Chapter III of the AI Act. Standalone high-risk AI systems – high-risk because they fall within an Annex III domain – must comply by 2 December 2027.[37] High-risk AI systems that are, or are a safety component of, a product covered by Annex I must comply by 2 August 2028. For high-risk AI systems already in use, the transitional deadlines of Article 111 AI Act apply: systems placed on the market or put into service before 2 August 2027 that are intended to be used by public authorities must comply by 2 August 2030, while other standalone high-risk systems must comply by 2 August 2027, only if, as from that date, those systems are subject to significant changes in their designs.

The obligations differ per role. The most extensive set of obligations applies to the provider, deployer, importer and distributor. Each carrying a more limited set. The relevant legal bases are Article 16 (provider), Articles 26 and 27 (deployer), Article 23 (importer) and Article 24 (distributor) of the AI Act.

A special case applies where the AI system is high-risk because it falls under Annex I Section B sectoral legislation – covering, among others, civil aviation security, motor vehicles and rail interoperability – that requires a third-party conformity assessment. In that case the AI Act does not impose new obligations beyond the sectoral framework. Obligations of Articles 8-17 are integrated into the relevant sectoral regulation (see Article 2(2)). More information on this can be found on the website of Algorithm Audit.[38]

Provider

The provider of a high-risk AI system carries the most extensive obligations. Pursuant to Article 16, the provider must satisfy the following 14 requirements:

1. **Risk management system –** Establish, implement, document and maintain a risk management system throughout the entire lifecycle of the high-risk AI system (Article 9).
2. **Data and data governance –** Ensure that training, validation and testing data sets are relevant, sufficiently representative and, to the best extent possible, free of errors (Article 10).
3. **Record-keeping and logging –** Retain automatically generated logs for at least six months, or as otherwise required by law (Article 12).
4. **Transparency and provision of information to deployers –** Provide clear instructions for use to deployers (Article 13).
5. **Human oversight –** Build in technical measures enabling effective human oversight and meaningful human intervention (Article 14).
6. **Accuracy, robustness and cybersecurity –** Design and develop the system to achieve an appropriate level of accuracy, robustness and cybersecurity throughout its lifecycle (Article 15).
7. **Quality management system –** Establish and maintain a quality management system, including post-market monitoring (Article 17).
8. **Technical documentation –** Prepare and maintain full technical documentation (Annex IV) before placing the system on the market (Article 11).
9. **Conformity assessment –** Carry out the applicable conformity assessment procedure – self-assessment or third-party notified body, depending on the system type (Article 43).
10. **CE marking –** Affix the CE marking to indicate conformity with the AI Act (Article 48).
11. **EU Declaration of Conformity –** Draw up and sign the declaration (Article 47).
12. **Registration –** Register the system in the EU database before placing it on the market (Article 49).

[37] Digital Omnibus on AI, European Commission (2026).
[38] Third party conformity assessment of high-risk AI systems, Algorithm Audit (2026).

13. **AI literacy –** Ensure that staff have an adequate level of AI literacy (Article 4).
14. **Post market monitoring -** Establish a monitoring system and report serious incidents (Art. 72–73)

Deployer

Pursuant to Articles 26 and 27, the deployer of a high-risk AI system must satisfy the following 9 requirements:

1. **Risk management system –** Monitor the operation of the system on the basis of the instructions for use and report identified risks or serious incidents to the provider, distributor and market surveillance authority without undue delay (Article 26(5)).
2. **Data and data governance –** Where the deployer controls the input data, ensure it is relevant and sufficiently representative in view of the intended purpose (Article 26(4)).
3. **Record-keeping and logging –** Keep the automatically generated logs for at least six months (Article 26(6)).
4. **Transparency and instructions for use –** Use the AI system in accordance with the instructions for use and take appropriate technical and organisational measures to ensure compliance (Article 26(1)).
5. **Human oversight –** Assign human oversight to persons with the necessary competence, training and authority (Article 26(2)).
6. **Workers –** Inform workers' representatives and affected workers before putting the system into service at the workplace (Article 26(7)).
7. **Transparency to affected persons –** Inform natural persons that they are subject to a high-risk AI system (Article 26(11)).
8. **Fundamental rights impact assessment –** Carry out a fundamental rights impact assessment before deployment. This applies only to public bodies and to certain private entities deploying systems listed in Annex III (Article 27). Whether it applies in a given case is determined in the Risk category questionnaire, as described in Chapter 6.
9. **AI literacy –** Ensure that staff have an adequate level of AI literacy (Article 4).

Importer

Pursuant to Article 23, the importer of a high-risk AI system must satisfy the following 4 requirements:

1. **Conformity assessment –** Verify, before placing the system on the market, that the provider has carried out the conformity assessment, drawn up the technical documentation and affixed the CE marking.
2. **Authoritative representative –** Verify that the provider has appointed an authorised representative, where required (only when providers are established in third countries and AI system falls within the high-risk category).
3. **CE marking and declaration of conformity –** Ensure the AI system bears the required CE marking and is accompanied by the EU Declaration of Conformity and instructions for use, of which a copy is maintained for ten years.
4. **Cooperation and notification –** Report any non-compliance or serious risk to the provider and competent authorities. Cooperate with national authorities and provide information upon request.

Distributor

Pursuant to Article 24, the distributor of a high-risk AI system must satisfy the following 3 requirements:

1. **CE marking and declaration of conformity –** Verify, before distributing the system, that it bears the required CE marking, is accompanied by the EU Declaration of Conformity and instructions for use, and that the provider and importer obligations have been met.
2. **Conformity –** Ensure that storage and transport conditions do not affect the system's conformity.
3. **Cooperation and notification –** Report any non-compliance or serious risk to the provider

and competent authorities. Cooperate with national authorities and provide information upon request.

7.3 Questionnaire Art. 50

The Article 50 questionnaire determines the transparency obligations that apply to the AI system based on its applications, role and the possible applicability of exceptions.

Article 50 distinguishes four sub-cases:

- **Article 50(1) –** AI systems interacting directly with natural persons.
- **Article 50(2) –** AI systems generating synthetic audio, image, video or text content.
- **Article 50(3) –** Emotion recognition and biometric categorisation systems.
- **Article 50(4) –** AI systems generating or manipulating deep fakes and AI-generated text published to inform the public on matters of public interest.

In Question 11 of the Risk category questionnaire, the user is asked whether either one of the scenarios listed above applies to the AI system. When the user selects 'It generates synthetic audio, image, video or text content' the answer to the follow-up question Q12 determines whether it entails non-consensual fake nude imagery, in which case the user is prompted with the prohibited outcome. In all other cases, selecting either one of the scenarios adds a 'generative and interactive' AI badge to the outcome screen. When this badge is activated, the user is prompted with the suggestion to start an additional questionnaire in the Obligations menu that will provide more detailed guidance for Article 50 of the AI Act. These questions will further determine whether Article 50 indeed applies to the AI system and what the exact obligations entail and are described below.

The deadline for the Article 50 transparency requirements is 2 August 2026. Only the deadline for the marking (watermarking) requirements under Article 50(2) is extended to 2 December 2026, given that a system was placed on the market before 2 August 2026. New systems must comply from 2 August 2026. The below questions are enriched by guidance and examples provided by the European Commission in the 'Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation' (hereafter: 'guidelines').[39]

Art. 50 Q1 – Does one of the following scenarios apply to the AI system?

When the user has started the detailed Article 50 questionnaire from the Obligations menu, it is prompted with Q1. Q1 determines which of the Article 50 sub-articles might apply to the AI system. Selecting one of the Article 50 scenarios first in Q1 directs the user to the Article 50 horizontal exception screen, before being routed to the follow-up questions per scenario. For each scenario that is selected in Q1, a follow-up question (Q2-Q6) determines whether the transparency obligations for that specific scenario apply. Selecting the answer option 'None of the above' in Q1 yields the 'No obligations under Article 50 apply' outcome.

Q1 consist of the following answer options:

- It interacts directly with natural persons → Article 50 horizontal exceptions, then Q2
  - Interaction entails the capability of a bidirectional exchange of information or actions between natural persons and the AI system with a genuine conversational or responsive character. It can occur in various forms (e.g., auditory, visual and physical). Natural persons should be able to provide input to the AI system, while the AI system should be capable of responding with contextual output. The interaction between

[39] Guidelines on transparency obligations for providers and deployers of certain AI systems, European Commission (2026).

the AI system and the natural persons must be direct. This typically involves real-time or near real-time interaction. AI agents are covered if they are capable of interacting with the persons instructing them or with other natural persons in the execution of the tasks (e.g., making bookings, managing correspondence, negotiating or concluding contracts, executing purchases, etc). Examples include: AI-enabled voice assistants, AI hotline for fraud, AI bots on social media.

> It generates or manipulates synthetic audio, image, video or text content → Article 50 horizontal exceptions, then Q3

Content generation refers to the generation of synthetic material by the AI system. It may include AI systems designed with a narrow intended purpose to produce specific outputs, but it may apply to AI systems that may serve multiple purposed, including GPAI systems and can produce a variety of types of content, as well as agentic AI systems, so long as they generate synthetic audio, image, video or text content. Content manipulation refers to already existing content (whether synthetic or not) that is altered by an AI system.

> It recognises emotions or categorises persons based on biometric data → Article 50 horizontal exceptions, then Q4

The purpose of the transparency obligation here is to ensure that natural persons are aware that they are exposed to emotion recognition and biometric categorisation systems that can be intrusive for their privacy. The obligation therefore applies regardless of whether the persons are exposed to such systems in real-time or those systems are operated ex post.

> It generates or manipulates image, audio or video content that resembles real people, objects, places or events (a deep fake) → Article 50 horizontal exceptions, then Q5

AI-generated or manipulated image, audio or video content appreciably resembles a subject if there is a high level of similarity between the deep fake content and the subject being simulated by the deep fake. The content does not need to be identical to the subject. It is sufficient for a deep fake to resemble someone or something that exists, can plausibly exist or could have plausibly existed. By contrast, simulated persons, objects, places, entities or events that, for example, defy the laws of nature or physics or depict lifeforms that are not commonly accepted in biology (e.g., humans flying without mechanical aids, dragons, or elephants driving cars) and have no potential to mislead are considered unrealistic and therefore fall outside this scope. Examples include: AI-generated audio involving voice cloning of a newspaper podcast's regular presenters and a guest interlocutor discussing some news events; AI-generated video of an individual resembling a politician, holding a speech in front of an audience. Examples that do not constitute deep fakes: AI-generated image of a scene involving a sphinx flying over the Eiffel Tower; AI voice replication for fictional characters (e.g., for audiobooks, games or animation) when there is no deception as to the identity of the narrators.

> It generates or manipulates text published to inform the public on matters of public interest → Article 50 horizontal exceptions, then Q6

> The AI-generated or manipulated text should be published. This means that the text should be accessible by an indeterminate, fairly large number of unrelated, potential readers simultaneously and/or successively,

whether or not against payment. By contrast, text is not considered published if access is restricted to specific individuals belonging to a closed, private group.

> The AI-generated or manipulated published text should aim to inform the public. This entails that the text should intend to communicate knowledge, opinions or facts.

> The text should inform the public 'on matters of public interest'. In general, such matters should be understood to cover those relevant to society at large, whether at a local, national, Union or international level, and meriting public debate or scrutiny. In that regard, texts should be considered to address public interest matters if they cover topics on politics and democratic processes, public administration and services, the administration of justice and law enforcement, the protection of fundamental rights, public security, public health, environmental protection, consumer safety, and any economic, financial, political, scientific, or cultural development that may be relevant subject of public debate. Matters that may be considered to be of public interest can evolve over time and across contexts.

Examples include: AI-generated summary of a human-authored article on a newspaper's website discussing a recent decision by a town council; AI-manipulated parts of a lifestyle-website article comparing the effects of various diets on a particular disease in middle-aged women; AI-manipulated corporate reports published on a listed company's website containing investor information; AI-generated message on a meteorological institute's social media profile warning citizens about stormy weather and related precautionary measures.

> None of the above → ends questionnaire

Art. 50 Article 50 horizontal exception – Does one of the following exceptions apply to your application?

The Article 50 exception question is shown directly after Q1, whenever at least one of the scenarios in Q1 is selected. It is asked before the sub-article specific exception questions (Q2-Q6), so that a horizontal exception to Article 50 is identified before the sub-article specific assessment starts.

The Article 50 horizontal exception question asks whether the AI system falls under one of the Article 50 exceptions:

> Deployers acting in the context of purely personal non-professional activity (guidelines paragraph 2.4.1)

Any activity through which natural persons gain an economic benefit on a regular basis or are otherwise involved in a professional, business, trade, occupational or freelance activity should be considered as a 'professional' activity. The specification of 'purely personal' is a qualifier of non-professional, meaning that the person should act in both a personal and a non-professional capacity to fall within the scope of the exception. Examples include: individuals using an AI system to generate deep fakes of the members of their household; individuals using an AI system in a personal capacity to generate text on various topics (e.g., students using AI to write their homework); individuals using an AI system in their personal capacity to generate deep fakes disseminated on social media.

> Research and development (guidelines paragraph 2.4.2)
  - This includes AI systems developed by students as part of their studies or by scientists affiliated with a university. In that case, the research must follow recognized ethical and professional standards for scientific research. For example, all participants must be informed that they are part of scientific research. The covers not only interactive AI systems, but also the outputs of generative AI systems (including deep fake content) used in the context of scientific research. However, if these systems are put into service or their outputs are being used for other purposes than the sole purpose of scientific research and development, this exception does not apply, and the relevant transparency obligations as stated by Article 50 of the AI Act would still need to be complied with.
> AI released under free and open-source licenses (guidelines paragraph 2.4.2)
  - Providers of free and open-source AI components that cover software, data, models, tools, services or processes for customisation and integration into an AI system and that do not constitute in themselves (in part of in whole) an AI system, are not subject to the transparency obligations in Article 50 of the AI act.
> None of the above

If a research exception is selected, the user is asked to describe the research (purpose, department, partners). When one of the three exceptions is selected, the questionnaire does not continue to the sub-article specific exception questions: the user is directed to the respective generative/interactive-ai-with-exception outcome, depending on the scenario or scenarios selected in Q1. When 'None of the above' is selected, the questionnaire continues with the sub-article specific exception question that corresponds to the scenario selected in Q1 (Q2-Q6).

To assist users, the following remark is provided:

When the AI system is used for operational processes or regular business operations, the user should select 'None of the above'.

Art. 50 Q2 – Your AI system interacts directly with natural persons. Which of the following scenarios apply to your AI system?

Q2 is shown when the user selected 'It interacts directly with natural persons' in Q1 and 'None of the above' in the Article 50 horizontal exception screen. Q2 is intended to identify whether an exception to Article 50(1) might apply according to considerations 41-49 of the guidelines:

> It is obvious to a well-informed, observant and alert natural person that it is interacting with an AI system
  - Consider the intended and reasonably foreseeable audience, not just the target audience. If children, elderly people, persons with disabilities or the general public could be exposed, the bar for 'obvious' is higher. Examples where this can apply: code-assistance tools used only by professional developers; internal staff assistants for AI-literate employees; medical-diagnosis support tools used only by trained professionals; single-player game characters where no other person can participate. It does not apply to realistic companion robots or avatars, or immersive VR/AR environments with human-like avatars, if the resemblance is high enough to create doubt, even for a professional audience.
> The AI system is authorised by law for law enforcement purposed
  - To fall within this scenario, the purpose of the AI system must be to detect, prevent, investigate or prosecute criminal offences.

It is not restricted only to be used by law enforcement authorities but may also cover AI systems used by other Union or national public authorities, so long as their use is authorised by law to detect, prevent, investigate or prosecute criminal offences, and subject to appropriate safeguards to protect the rights and freedoms of third parties. If the AI system is available to the public and offers a functionality to individuals to report criminal offences it is not included.

> None of the above

If the user selects one of the two answers other than 'None of the above' the exception to the obligations under Article 50(1) might apply. The user is then redirected to the Art. 50 outcome screen with the outcome interactive-ai-with-exception. A 'None of the above' routes the user directly to the Interactive AI outcome where the obligations and guidelines under Article 50(1) of the AI Act are shown.

Art. 50 Q3 – Your AI system generates or manipulates synthetic audio, image, video or text content. Which of the following scenarios apply to your AI system?

Q3 is shown when the user selected 'It generates or manipulates synthetic audio, image, video or text content' in Q1 and 'None of the above' in the Article 50 exception screen. Q3 is intended to identify whether an exception to Article 50(2) might apply according to considerations 89-93 of the guidelines:

> The AI system performs an assistive function for standard editing

Standard editing should be understood as the process of preparing existing content for publication or distribution (e.g., small edits to improve readability and grammar, quality and format) and does not involve generating new content. Standard editing aims at ensuring that content is, among others, free from obvious technical or grammatical linguistic errors, is aligned with applicable lay-out, presentation, formatting or accessibility requirements, or in conformity with sectoral practices. Editing goes beyond standard editing if the content is changed in a material way (substantive modifications, structural changes etc.) that affect its meaning, style or intent.

> The AI system does not substantially alter the input data provided by the deployer or the semantics thereof

An alteration should be considered substantial if the input data or its semantics have been manipulated significantly by the AI system during its output generation process. This must be based on an assessment of relevant factors, such as the format, media content type, style and changes in the content that affect its meaning, style or intent. Whether that is the case requires a case-specific assessment.

> The AI system is authorised by law to generate or manipulate synthetic content to detect, prevent, investigate or prosecute criminal offence

To fall within this scenario, the purpose of the AI system must be to detect, prevent, investigate or prosecute criminal offences. It is not restricted only to be used by law enforcement authorities but may also cover AI systems used by other Union or national public authorities, so long as their use is authorised by law to detect, prevent, investigate or prosecute criminal offences, and subject to appropriate safeguards to protect the rights and freedoms of third parties. If the AI system is available to the public and offers a functionality to individuals to report criminal offences it is not included.

> None of the above

If the user selects one of the three answers other than 'None of the above' the exception to the obligations

under Article 50(2) might apply. The user is then redirected to the Art. 50 outcome screen with the outcome generative-ai-with-exception. A 'None of the above' routes the user directly to the Generative AI outcome where the obligations and guidelines under Article 50(2) of the AI Act are shown.

Art. 50 Q4 – Your AI system recognises emotions or categorises persons based on biometric data. Is it used for the purpose of law enforcement?

Q4 is shown when the user selected 'It recognises emotions or categorises persons based on biometric data' in Q1 and 'None of the above' in the Article 50 horizontal exception screen. Q4 is intended to identify whether an exception to Article 50(3) might be made based on the purpose of law enforcement according to the consideration 109 of the guidelines. If the user chooses 'Yes' the exception to the obligations under Article 50(3) might apply. The user is then redirected to the Art. 50 outcome screen with the outcome emotion-recognition/biometric-categorisation-with-exception. A 'No' routes the user directly to the emotion recognition/biometric categorisation outcome where the obligations and guidelines under Article 50(3) of the AI Act are shown.

To assist users, the following remark is provided:

When the AI system uses emotion recognition or biometric categorisation that is permitted by law to detect, prevent or investigate criminal offences subject to appropriate safeguard for the rights and freedoms of third parties and in accordance with Union law, choose 'Yes'.

It is not restricted only to be used by law enforcement authorities but may also cover AI systems used by other Union or national public authorities, so long as their use is authorised by law to detect, prevent, investigate, or prosecute criminal offences, and subject to appropriate safeguards to protect the rights and freedoms of third parties. If the AI system is available to the public and offers a functionality to individuals to report criminal offences, choose 'No'.

Art. 50 Q5 – Your AI system generates or manipulates image, audio or video content that resembles real people, objects, places or events (a deep fake). Which of the following scenarios apply to your AI system?

Q5 is shown when the user selected 'It generates or manipulates image, audio or video content that resembles real people, objects, places or events (a deep fake)' in Q1 and 'None of the above' in the Article 50 horizontal exception screen. Q5 is intended to identify whether an exception to Article 50(4) might apply according to considerations 119-125 of the guidelines:

- The deep fake forms a part of evidently artistic, creative, satirical, fictional or analogous work
  - Artistic works are to be understood as works that have been created for the purpose of art, including music, cinematographic works and visual arts. E.g., movies (or movie trailers) featuring AI-generated or manipulated deep fakes of de-aged, existing actors or digital replicas of dead actors or other persons, available in a cinema or on a streaming platform.
  - Creative works are to be understood as works that involve creative choices. Works that are mainly motivated by functional or technical considerations cannot be regarded as creative. E.g., AI-generated music in any kind of genre resembling the individual style of existing artists.
  - Satirical works are to be understood as works that are intended to criticise society, politics, business or public figures through the use of humoristic

techniques (including irony, sarcasm, mockery, pastiche etc.) E.g., AI-manipulated image of an existing politician placing him in a scene clearly meant to criticise in a humorous manner certain policy decisions taken or supported by that person.

- Fictional works are to be understood as cultural works that involve persons, objects, places, entities or events, in an imaginary, but verisimilitude setting. E.g., AI-generated gaming imagery involving deep fake simulations of real, existing persons.
- Analogous works are to be understood as works sharing core traits or similar expressive or functional characteristics with the above categories but not fitting neatly into one.
- The fact that a deep fake falls within one (or more) of the above categories should be evident to the natural persons exposed to it.

> The deep fake is authorised by law to detect, prevent, investigate or prosecute criminal offences

To fall within this scenario, the purpose of the deep fake must be to detect, prevent, investigate or prosecute criminal offences. It is not restricted only to be used by law enforcement authorities but may also cover AI systems used by other Union or national public authorities, so long as their use is authorised by law to detect, prevent, investigate, or prosecute criminal offences, and subject to appropriate safeguards to protect the rights and freedoms of third parties. If the AI system is available to the public and offers a functionality to individuals to report criminal offences it is not included.

> None of the above

If the user selects 'The deep fake forms a part of evidently artistic, creative, satirical, fictional or analogous work' a badge is added to the outcome that indicates that the transparency obligation for this content is limited to the disclosure of the deep fake in an appropriate manner that does not hamper the display or enjoyment of the work. It does however fall within Article 5(4) and therefore the user is routed directly to the Deep fake outcome where the obligations and guidelines under Article 50(4) of the AI Act are shown.

If the user selects 'The deep fake is authorised by law to detect, prevent, investigate or prosecute criminal offences', an exception to Article 50(4) applies and the user is directed to the deepfake-with-exception outcome screen.

A 'None of the above' routes the user directly to the Deep fake outcome where the obligations and guidelines under Article 50(4) of the AI Act are shown.

Art. 50 Q6 – Your AI system generates or manipulates text published to inform the public on matters of public interest. Which of the following scenarios apply to your AI system?

Q6 is shown when the user selected 'It generates or manipulates text published to inform the public on matters of public interest' in Q1 and 'None of the above' in the Article 50 horizontal exception screen. Q6 is intended to identify whether an exception to Article 50(4) might apply according to the considerations 133-139 of the guidelines:

> The text is under human review or editorial control and editorial responsibility

Two conditions must be met for this scenario: (i) the AI generated or manipulated text must have undergone human review or editorial control and (ii) a legal or natural person must hold editorial responsibility for the publication.

- Human review refers to the deliberate examination of the substance of the content by one or more natural persons possessing relevant knowledge and professional judgement pertaining to the subject matter under scrutiny (e.g., academic peer review or professional validation chains). Fact-checking the accuracy of the content is a minimum requirement that should be part of that review. Editorial control refers to the control exercised in practice by a responsible editorial entity (e.g., an editor-in-chief) over the content having the authority to approve, alter or reject the substance of the text based on substantive grounds (incl. fact-checking of information and ensuring the trustworthiness of sources). Superficial, solely formal or procedural checks (e.g., spell-checking or grammatical correction), the mere existence of an editorial policy, automated review processes or cursory editorial approval without substantive engagement by the human reviewer or the editorial entity, cannot fulfil the conditions for human review or editorial control.
- Editorial responsibility entails that said person must hold the ultimate legal responsibility over the publication of the content, including the human review or editorial control (e.g., an individual, editorial board, or the publishing company). To ensure public accountability and trust, and in line with existing media professional standards, the identity and contact details of the legal person, the natural person or the function with editorial responsibility should be made publicly available on an easily findable location (if not yet otherwise available). This can happen online through, e.g., a website's terms and conditions or other user-facing legal information. Offline, such information can be included in a publication's colophon or edition notice.

  Examples include: An AI-manipulated article in a newspaper or AI-generated summary of an article that has been subject to editorial control of the respective editor-in-chief with editorial responsibility held by the legal person that publishes the newspaper; An AI-manipulated academic blog which has undergone internal peer review and where the respective research centre managing the blog holds editorial responsibility; AI-generated public safety warnings approved by a public official before being distributed to citizens, under the responsibility of the relevant public agency for civil protection; AI-generated sustainability reports published on a listed company's website having undergone human review by professionals in relevant functions (e.g., compliance); AI-supported translation of a human-written article whereby the translation has undergone human review.
- The use of the AI system for generating text published with the purpose of informing the public of on matters of public interest without human review is authorised by law to detect, prevent, investigate or prosecute criminal offences

  To fall within this scenario, the purpose of the generated text must be to detect, prevent, investigate, or prosecute criminal offences. It is not restricted only to be used by law enforcement authorities but may also cover AI systems used by other Union or national public authorities, so long as their use is authorised by law to detect, prevent, investigate or prosecute criminal offences,

and subject to appropriate safeguards to protect the rights and freedoms of third parties. If the AI system is available to the public and offers a functionality to individuals to report criminal offences it is not included.

> None of the above

If the user selects 'The text is under human review or editorial control and editorial responsibility' or 'The use of the AI system for generating text published with the purpose of informing the public of on matters of public interest without human review is authorised by law to detect, prevent, investigate or prosecute criminal offences', an exception to Article 50(4) might apply and the user is directed to the public-interest-text-with-exception outcome screen.

A 'None of the above' routes the user directly to the public interest text outcome where the obligations and guidelines under Article 50(4) of the AI Act are shown.

7.4 Obligations for generative and interactive AI

The transparency obligations differ per role and per sub-case. The information must always be provided in a clear and distinguishable manner, at the latest at the time of the first interaction or exposure (Article 50(5)). The sub-sections below set out the obligations for providers and deployers.

Provider

Interacting directly with natural persons (Article 50(1))

The provider must:

> **Clearly informed –** Design and develop the AI system so that natural persons are clearly informed that they are interacting with an AI system, unless this is obvious from the perspective of a reasonably well-informed natural person, taking into account the circumstances and context of use.
> **First interaction –** Embed the notification in the design of the system and provide it at the latest at the time of the first interaction (Article 50(5));
> **Accessible –** Ensure the information is clear, distinguishable, accessible – including for vulnerable groups such as children, the elderly and persons with disabilities – and given in a context-appropriate manner.

Generating synthetic content (Article 50(2))

The provider must:

> **Machine-readable marking –** Mark the outputs in a machine-readable format, for example through watermarks, metadata, cryptographic provenance methods or fingerprints.
> **Human-readable –** Make means available to detect that the output is AI-generated or manipulated, with human-readable results.

Both the marking and detection elements must be present. Marking alone is insufficient. The technical solution must be effective, reliable, robust and interoperable.

Emotion recognition and biometric categorisation (Article 50(3))

Obligations apply to deployers only and are set out below.

Deep fakes and AI-generated public-interest text (Article 50(4))

Obligations apply to deployers only and are set out below.

Deployer

Interacting directly with natural persons (Article 50(1))

The deployer must:

> **Verify –** Verify that the provider has implemented the disclosure of interactive AI, for example an opening line from a chatbot, a clear "AI assistant" label, or an equivalent signal.
> **Does not undermine or remove –** Ensure that

its own configuration, branding or integration of the system does not undermine or remove the provider's disclosure. Where the deployer fine-tunes, rebrands or substantially modifies the system, Article 25 may convert the deployer into a provider, importing the full Article 50(1) provider obligations.

Generating synthetic content (Article 50(2))

Obligations apply to providers only and are set out above.

Emotion recognition and biometric categorisation (Article 50(3))

The deployer must:

- **Inform –** Inform all natural persons exposed to the system that it is being operated or used.
- **Accessible –** Provide the information in a clear, distinguishable and accessible manner (Article 50(5)).
- **Standardised icons –** The information may be given in writing, via standardised icons, orally, electronically, or in any combination, depending on the deployment context.

Timing: Provide the information to users at the latest at the time of the first exposure to the system.

Deep fakes and public-interest text (Article 50(4))

The deployer must:

- **Deep fake –** Disclose that content constituting a deep fake (as defined in Article 3(60)) has been artificially generated or manipulated.
- **Public interest –** Disclose that AI-generated or manipulated text published with the purpose of informing the public on matters of public interest has been artificially generated or manipulated.

Timing: This information must be provided to users at the latest at the time of the first exposure to the system.

Exceptions: The disclosure obligation does not apply where the use is authorised by law to detect, prevent, investigate or prosecute criminal offences; where the content forms part of an evidently artistic, creative, satirical, fictional or analogous work or programme – in which case the obligation is limited to disclosing the existence of generated or manipulated content in a manner that does not hamper the display or enjoyment of the work; or where the AI-generated content has undergone a process of human review or editorial control and a natural or legal person holds editorial responsibility for the publication.

For Article 50(1) and Article 50(2), the guidelines provide further detail on acceptable and (in) sufficient disclosure techniques and on the quality requirements for marking and detection solutions.

Flowchart Obligations

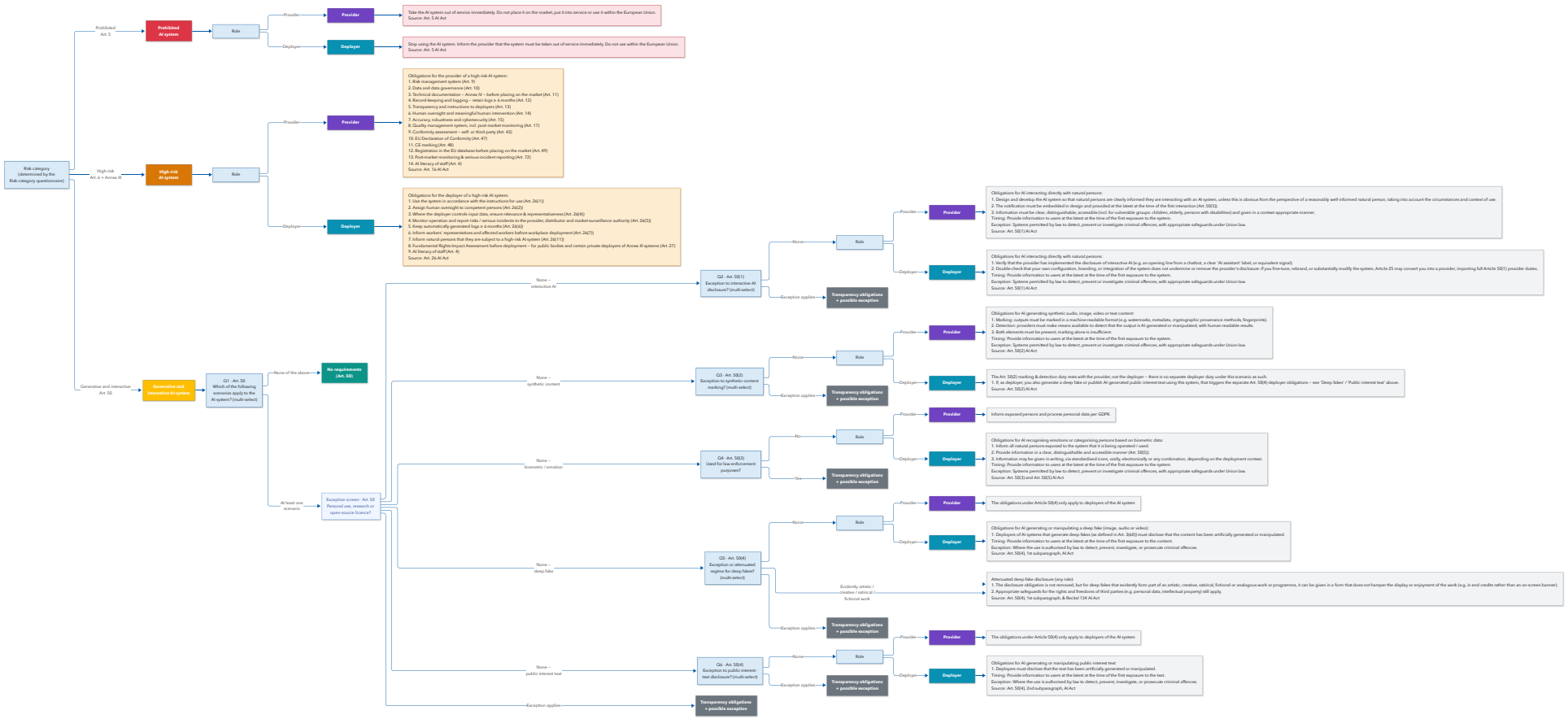


Figure 29 - Flowchart determining the obligations that apply to the AI system

About Algorithm Audit

Algorithm Audit is a European knowledge platform for AI bias testing and normative AI standards. The goals of the NGO are four-fold:

Knowledge platform	Bringing together experts and knowledge to foster the collective learning process on the responsible use of algorithms, see for instance our AI Policy Observatory and position papers
Normative advice commissions	Forming diverse, independent normative advice commissions that advise on ethical issues emerging in real world use cases, resulting over time in algoprudence
Technical tools	Implementing and testing technical tools for bias detection and mitigation, e.g, bias detection tool, synthetic data generation and sociotechnical evaluation of generative AI
Project work	Support for specific questions from public and private sector organisations regarding responsible use of AI

Structural partners of Algorithm Audit

SIDN Fund

The SIDN Fund stands for a strong internet for all. The Fund invests in bold projects with added societal value that contribute to a strong internet, strong internet users, or that focus on the internet's significance for public values and society.

European AI&Society Fund

The European AI&Society Fund supports organisations from entire Europe that shape human and society centered AI policy. The Fund is a collaboration of 14 European and American philantropic organisations.

Public knowledge building for *responsible* AI

www.algorithmaudit.eu

www.github.com/NGO-Algorithm-Audit

info@algorithmaudit.eu

Parkstraat 22, 2514 JK Den Haag

Algorithm Audit

Stichting Algorithm Audit is registered as a non-profit organisation at the Dutch Chambre of Commerce under license number 83979212