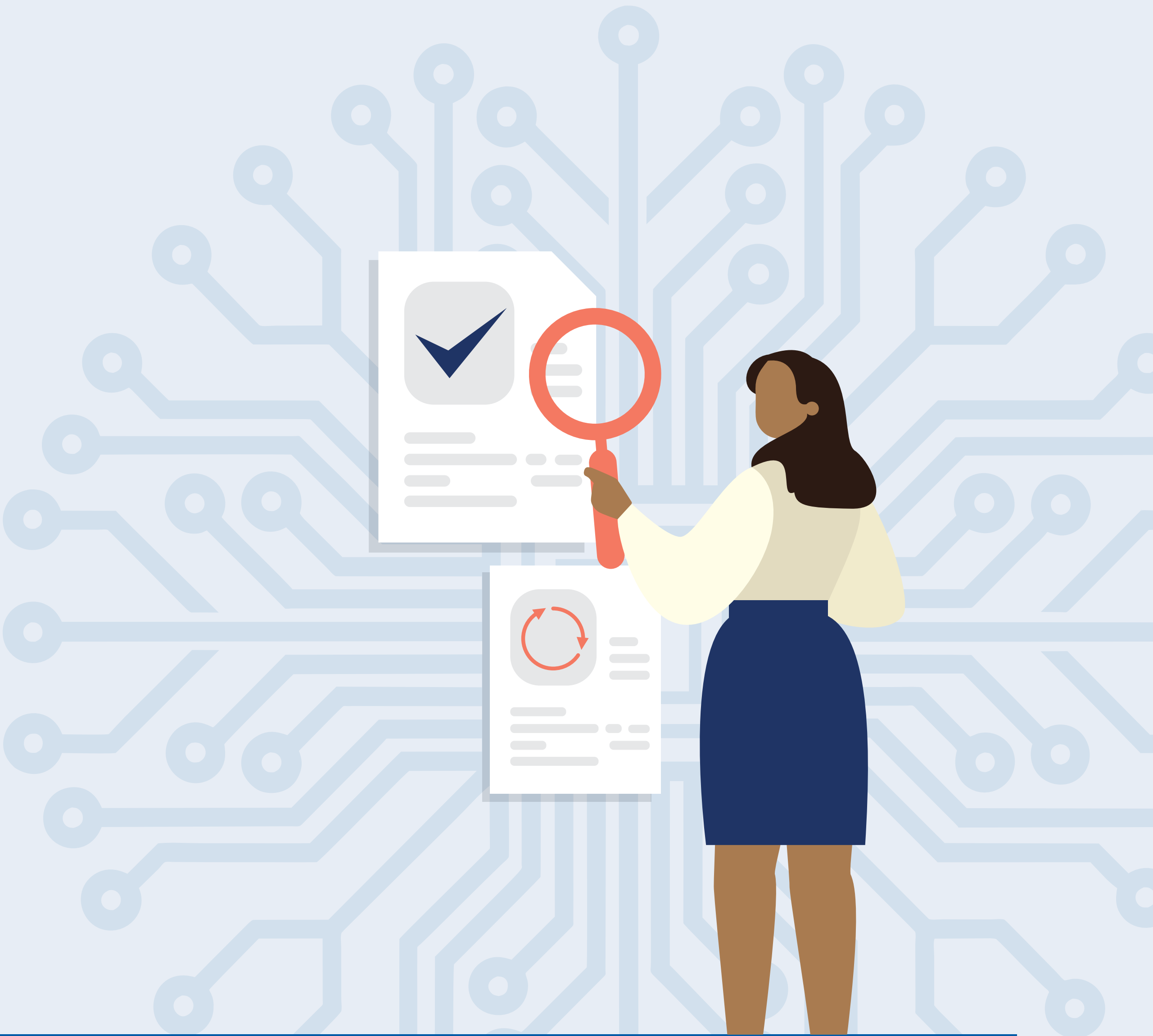


Implementation of the AI Act

Definition of an AI system

February 2025

1. Introduction

The AI Act sets requirements for artificial intelligence (AI) in the European Union (EU). The product safety legislation oversees the responsible development, deployment and use of AI by public and private organisations. This protects the safety, health and fundamental rights of EU citizens. However, the implementation of the AI Act raises difficult issues, such as which algorithmic applications fall within the scope of the regulation.

Neither in the societal debate around this technology, nor within the academic and technical world, there has been a fixed, shared definition the term AI in the past 50 years. Most people use intuitive, unwritten definitions when discussing AI. What is seen as AI evolves with the technological cutting edge: as soon as generally accessible software can perform complex tasks that were previously reserved for 'AI', it is soon no longer seen as AI.[1]

However, with the advent of the AI Act, AI is captured in a legally binding definition. The EU's goal with this definition is to distinguish AI systems from simpler traditional software systems or programming approaches, thereby providing legal certainty, foster wide acceptance and ensure a future-proof definition.[2] The definition as used by the European legislator is not new: it follows the definition of AI as developed by the Organisation for Economic Co-operation and Development (OECD).

Based on this legal definition, organizations must start implementing the AI Act. This turns out to be complicated. For example, lawyers often have little practical experience with the technologies underlying AI and technicians are inexperienced with legal definitions. In addition, the definition of an AI system contains terms that do not all have the same weight. For the implementation of the regulation, it is therefore necessary to build bridges between these different worlds, taking into account not only completeness but also pragmatism.

This white paper makes a first step in this direction. We analyse the key elements from the definition of an AI system from both a legal and statistical perspective, exploring the scope of the AI Regulation. The European Commission's (EC) guidelines on the definition of an AI system have been included in this analysis.[3]

In addition, the relationship between AI systems and high impact algorithms is discussed. 'Algorithms'[4] is a term that has been used for some time by the Dutch government to refer to a broader category of automated systems, which also includes AI.[5] Algorithms can have significant impact on data subjects[6], even when they are not AI systems. These algorithms with an increased risk of impacting data subjects are referred to as high-impact algorithms.

1 Facial recognition and chess computers have long been seen as the ultimate example of AI, while the applications are now integrated into everyday life and are no longer referred to as such. This phenomenon has been described by Pamela McCorduck as "the AI effect".

2 See recital 12 AI Act.

3 'Guidelines on the defin[i]tion of an artificial intelligence system established by AI Act', European Commission (2025).

4 'Algorithm' as defined by the Dutch Court of Audit (2021): "A set of rules and instructions that a computer automatically follows when making calculations to solve a problem or answer a question".

5 Figure 3 Guideline Algorithm Register of the Dutch Ministry of the Interior and Kingdom Relations. The fact that high-impact algorithms are an important category of algorithms is evident from *Parliamentary Papers II 2024-25 2025D00512*.

6 For example, in 2023, after research by Investico, NOS op 3 and De Groene Amsterdammer, among others, it came to light that DUO used an algorithm for risk-based selection during control on the exchange for the exchange. After an investigation by PwC and the Algorithm Audit Foundation, it turned out that there was (indirect) discrimination. The Minister of Education, Culture and Science has announced a recovery operation, in which more than ten thousand students will be compensated.

A schematic overview of the different types of algorithms and the relationship between these concepts can be found in Figure 1.

We introduce three questions that users can use to determine whether an algorithmic application is an AI system. On the basis of a maximum of four additional questions, it can be determined whether it is a high-impact algorithm. These dynamic questionnaires are available online under the name AI Act Implementation Tool.[7] This tool is published under the EUPL-1.2 license.[8]

Identification of high-risk and prohibited AI systems are discussed in a separate white paper.[9]

This white paper analyses the seven characteristics of the definition of an AI system (sec. 2). Special attention is paid to the concept of inference (sec. 3) and autonomy (sec. 4). In addition, the relationship between AI systems and the concept of high-impact algorithms is discussed (sec. 5). The paper concludes with the dynamic questionnaires that can be used to identify AI systems and high-impact algorithms (sec. 6).

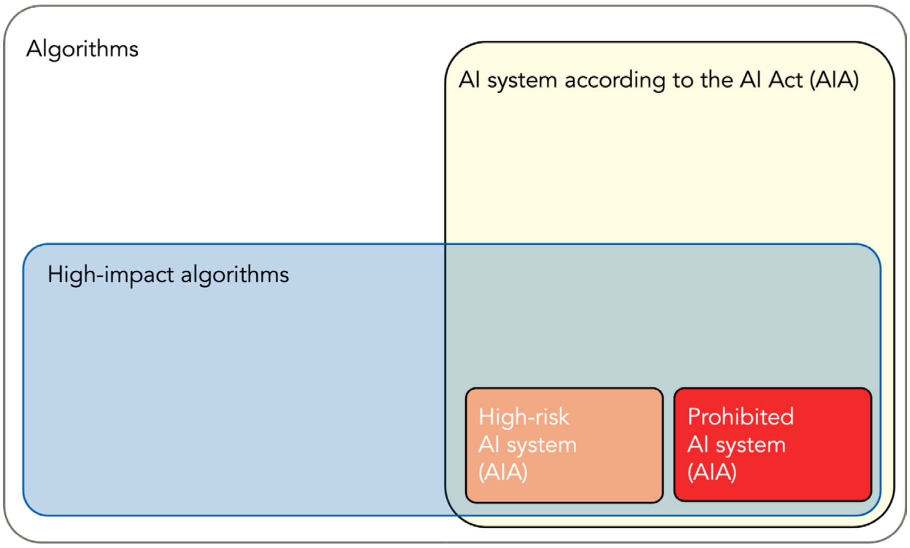


Figure 1 - Overview of the types of algorithms discussed in this white paper. High-risk and prohibited AI systems are discussed in a separate white paper.[4]

7 https://algorithmaudit.eu/technical-tools/implementation-tool/#tool
8 https://github.com/NGO-Algorithm-Audit/AI-Act-Implementation-Tool?tab=EUPL-1.2-1-ov-file
9 https://algorithmaudit.eu/technical-tools/implementation-tool/#documentation-high-risk

Box 1

Disclaimer regarding AI Act compliance

This document is an interpretation of the legal text of the AI Act and additional guidelines as published by the European Commission (EC) by Algorithm Audit. No rights can be derived from this analysis. As noted in paragraph (7) of the EC's issued guidelines, the Court of Justice of the European Union (CJEU) ultimately decides what the correct interpretation of an AI system is.

2. AI system definition

The definition of an AI system is introduced in Article 3(1) of the AI Act. Only systems that meet this definition are in scope of the Act.

Article 3(1) of the AI Act 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."

We analyse and interpret the above seven highlighted concepts. The main sources for this interpretation are recital 12 of the preamble to the Act and the guidelines as published by the EC (hereafter: guidelines).[10]

Recital 12 consists of 13 phrases which are included for reference in Appendix A. Recitals provide insight into the intentions of the EU legislature in the process of drafting the legal text and thus provide an explanation of how the concepts in the law should be interpreted.

The guidelines have been published to provide additional explanations on the definition of an AI system. In the analysis of the definition of an AI system, following in sections 2-4, specific passages from these guidelines are referred to continuously. In addition to an interpretation of the above seven coloured concepts that form the definition of an AI system, the guidelines also introduce a number of controversial exceptions. These exceptions obscure the interpretation of the definition instead of providing clarity about it. A substantive explanation of this criticism can be found in Box 2.

This white paper also refers to the OECD Memorandum on the definition of an AI system (hereafter: the OECD Memorandum).[11] This memorandum, including previous draft versions, was used during negotiations on the AI Act, to arrive at the definition of an AI system in the legal text. In this light, recital 12 explicitly mentions the EU's desire to *"be closely aligned with the work of international organisations working on AI to ensure legal certainty, facilitate international convergence and wide acceptance"*.

We conclude each analysis of the above seven coloured concepts with an assessment of the extent to which this concept can serve as a criterion to distinguish AI systems from algorithms. In view of the importance of the concepts of inference and autonomy in the definition of an AI system, they are analysed independently in 3. Inference and 4. Autonomy.

Building on this analysis, a dynamic questionnaire will be introduced that can be used to identify AI systems within three questions. See section 6.1-6.3 in 6. Dynamic questionnaire

[10] Supra note 3.

[11] Explanatory Memorandum on the Updated OECD definition of an AI system (2024).

Box 2

Exceptions in the guidelines create ambiguity in interpretating the AI system definition

The European Commission has published guidelines on how to interpret the definition of an AI system as stated in the AI Act.[12] The guidelines introduce exceptions for algorithmic systems that do not qualify as AI systems based on arguments that do not align with the legislative text. This is remarkable, as guidelines are intended to clarify the interpretation of the legislative text rather than introduce additional provisions. Consequently, the guidelines blur the interpretation of the legislative text instead of providing clarity.

Specifically, the exceptions for systems for improving mathematical optimization (paragraphs (42)-(45)) and simple prediction systems (paragraphs (49)-(51)) cause issues.

For example, paragraph (49) states that machine learning systems that make use of a *"basic statistical learning rule"* are not AI systems. However, in the definition of an AI system in article 3 of the AI Act and in its explanation in recital 12, the complexity of a system is not mentioned as a determining factor for qualifying as an AI system.[13] The exception creates confusion: when is a statistical learning rule 'basic' enough to fall under this exception?

Paragraph (42) states that systems used for mathematical optimization do not qualify as AI systems. However, according to the definition in the legislative text, the application does not determine whether a system is an AI system. This paragraph explains that established methods, such as linear and logistic regression, are not AI systems because *"while those models have the capacity to infer, they do not transcend 'basic data processing'"*. This passage directly contradicts recital 12 of the AI Act, which states that: *"The capacity of an AI system to infer transcends basic data processing by enabling learning, reasoning or modelling"*. Both the claim that mathematical optimization falls outside the scope of the definition and the explanation for this are in conflict with the legislative text.

Another inconsistency arises in the interpretation of the term 'adaptiveness.' Paragraphs (22)-(23) of the guidelines explain that adaptiveness is not a strict requirement to meet the definition of an AI system. However, later in paragraph (48), the guidelines state that heuristics do not qualify as AI systems due to a lack of 'adaptability'.

With the introduced exceptions in these guidelines, the Commission appears to be narrowing the politically negotiated definition of an AI system between the Parliament and Council, and imposing its own interpretation of the Act. From a democratic perspective, the European Commission seems to be overstepping its mandate. Given the tensions between the guidelines and the AI Act, it is important to note that the guidelines have a subordinate legal status compared to the legislative text in the hierarchy of regulatory instruments. Until case law from the Court of Justice of the European Union becomes available, Algorithm Audit advises organizations,

---

[12] Supra note 3.

[13] Recital 12 does mention that the definition itself should include the features that distinguish AI systems from simpler traditional software systems. Therefore, it must be assumed that a system that meets the characteristics in the definition is not a simple traditional software system.

in line with the position of the Dutch Data Protection Authority (Autoriteit Persoonsgegevens), to err on the side of caution when determining whether AI systems fall within the scope of the AI Act.[14]

Finally, we note that the developments mentioned above contribute to an attempt to narrow the scope of the AI Act. This time, not through a discussion on the scope of the risk classification of AI systems – since only high-risk AI systems are required to comply with mandatory control measures – but through the question of whether algorithmic systems even fall under the definition of an AI system. In this way, the scope of the AI Act is being attempted to be narrowed both through the route of AI system identification and risk classification.

[14] Report on AI & Algorithm Risks in the Netherlands, Winter 2024/2025 (Edition 4, February 2025), Directorate for the Coordination of Algorithms (DCA) – Dutch Data Protection Authority (DPA).

2.1 Interpretation of the definition of an AI system using recital 12

Recital 12 contains a number of phrases that help to frame the interpretation of the definition of an AI-system:

i) *"The definition [of an AI system] should be based on key characteristics of AI systems that distinguish it from simpler traditional software systems or programming approaches"*;
ii) *"[The definition of an AI system] should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations."* – see recital 12 sentence 2.

From phrase i) follows the lens through which we interpret the definition of an AI system: the characteristics in the definition must make it possible to distinguish between AI systems and other software systems. The sentence is also a lower limit with which the legislator indicates that the scope of the definition of an AI system does not cover all programming approaches. 'Simple traditional software systems' could be understood to mean simple data processing in Excel or SQL. Although these programming approaches can also carry out more advanced data processing, which may well involve an AI system. Phrase i) is therefore not relevant for determining whether or not an application is an AI system. This is confirmed by paragraph (26) of the guidelines.

Phrase ii) refers to rule-based algorithms where the rules are created by natural persons. An example of a rule is 'if age <65 years, then no right to a senior discount'. If the variable 'age' and the threshold of '65 years' were set solely by natural persons to perform the automatic action of determining the right to a discount, the rule-based algorithm is not an AI system. This is even the case when this algorithm is used for impactful purposes, such as risk profiling. Phrase ii) has the ability to differentiate AI systems from algorithms. This characteristic is therefore included as an answer option in the third question of the dynamic questionnaire. See 6.3 Q3 – Is the application automation of human-defined rules? The guidelines do not provide a specific explanation of the relationship between rule-based algorithms and AI systems.

2.2 Machine-based system

Recital 12 states that '*machine-based system*' has the following meaning from the AI system definition: *"The term 'machine-based' refers to the fact that AI systems run on machines."* – see recital 12, sentence 7.

Since virtually all modern software systems or programming approaches use a machine, be it a computer, server or virtual machine (VM)[15], virtually all software systems and algorithms meet this requirement.

We conclude that the 'machine-based system' requirement is not a characteristic that distinguishes AI systems from algorithms, because all modern software systems or programming approaches are machine-based. This is confirmed by paragraphs (11)-(13) of the guidelines.

2.3 Varying levels of autonomy

Recital 12 states that '*varying levels of autonomy*' from the AI system definition has the following meaning:
"AI systems are designed to 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." – see recital 12, sentence 12.

There must therefore be some degree of autonomy. That is why we see 'autonomy' as a characteristic that distinguishes AI systems from algorithms. In 4. Autonomy the meaning and interpretation of autonomy is discussed in more detail.

2.4 May exhibit adaptiveness

Recital 12 states that "*may exhibit adaptiveness after deployment*" has the following meaning from the AI system definition:
"The adaptiveness that an AI system could exhibit after deployment, refers to self-learning capabilities, allowing the system to change while in use"

The use of the verbs may and could, lead to the conclusion that adaptability of an AI system is not a requirement. This is confirmed by paragraphs (22)-(23) of the guidelines. The OECD also sees adaptability as optional after its deployment, in the memorandum it also explicitly mentions a system that has been learned once from data as an AI system.[16] Many AI systems currently in use do not show adaptability after deploying them. Face recognition software, which the AI Act refers to in various places, is an example where model parameters are generally not updated during use but only prior to a software release. In short, even AI systems that do not show adaptability during use can still be an AI system, if the other conditions are met.

We conclude that 'adaptability' is not a requirement for the AI system definition. Therefore, it is not a characteristic that distinguishes AI systems from algorithms.

2.5 Explicit or implicit objectives

Recital 12 states that "*for explicit or implicit objectives*" from the AI system definition has the following meaning:
"The reference to explicit or implicit objectives underscores that AI systems can operate according to explicit defined objectives or to implicit objectives. The objectives of the AI system may be different from the intended purpose of the AI system in a specific context." – see recital 12 sentence 8.

[15] A VM refers to a microprocessor that runs algorithms on a PC, laptop, or in a cloud environment. See also 3.32 of ISO/IEC 13522-6:1998 Information technology — Coding of multimedia and hypermedia information

[16] Supra note 11

An application always pursues a goal, which can be defined either explicitly or implicitly. The reason this element is included in the definition is to express that an explicit objective is not a requirement for an AI system.[17] For example, with reinforcement learning, AI systems can derive objectives themselves, which are not explicitly formulated but are implicit in the AI system. This is also the case with Large Language Models (LLMs) such as ChatGPT and other applications of generative AI. This picture is confirmed by paragraph (24) of the guidelines. It follows from paragraph (25) that the 'intended purpose' refers not only to internal operations carried out by a system, but also to the external context in which the system is applied.

'Explicit or implicit objectives' is not a characteristic that distinguishes AI systems from algorithms.

2.6 Infers, from the input it receives, how to generate outputs

Recital 12 states that "*infers, from the input it receives, how to generate outputs*" has the following meaning from the AI system definition:
"A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to derive models or algorithms, or both, from inputs or data." – see recital 12 sentence 3-4.

Recital 12 mentions 'inference' explicitly as a key characteristic. We conclude that the capability to infer is the most important element of the definition to distinguish AI systems from other algorithms. This is confirmed in paragraph 26 of the guidelines, which call inference an "*indispensable condition that distinguishes AI systems from other types of systems*".

In 3. Inference the meaning and interpretations of inference is discussed.

2.7 Predictions, content, recommendations or decisions

Recital 12 states that "*predictions, content, recommendations or decisions*" from the AI system definition has the following meaning:
"[…] outputs generated by the AI system reflect different functions performed by AI systems and include predictions, content, recommendations or decisions." – see recital 12 sentence 10.

This passage is related to the inference of output from input. An analysis of 'inference' follows in 3. Inference. With regard to "*predictions, content, recommendations or decisions*" these are different forms of output that are derived:

1. **Predictions:** This includes estimated scores, rankings, probabilities, labels, and classifications. This does not necessarily have to be a prediction about the future, a prediction can also relate to a data point that has not been observed before. The statistical term 'estimator' is also a prediction in this context.
2. **Content:** This includes generated text, images, and speech, for example created through generative AI.
3. **Recommendations:** This includes recommendation systems, such as personalized timelines on social media platforms, search engine results, and online advertising. This category also includes recommendation of actions, such as a recommendation for additional checks that follow an assigned risk score for unlawful use of a social facility, or a car that recommends shifting to a different gear.[18] Scores or classifications which are tied to a fixed action or procedural step can also be seen as recommendations. For example: an assigned risk score in transaction monitoring within banks, on the basis of which

[17] Supra note 11

[18] Supra note 11

a work instruction prescribes that additional research must be carried out. Recommendations often follow from a prediction.

4. **Decisions:** These include decisions in the broadest sense of the word, such as the decision to perform an action, for example a car that automatically brakes for a pedestrian[19], the choice to carry out an investigation, the determination of someone's identity (verification) or a formal decision by a governmental body as defined in national public administration law (for example Awb art.1:3 in the Netherlands).[20] For the public sector, it is important to note that algorithmic output used in the preparatory phase of a decision should also be considered as part of the entire decision-making process and should therefore also comply with the general principles of good administration (Abbb in the Netherland), such as the duty of care, the duty to give reasons and the principle of fair play.[21] When the output is a recommendation or decision, the concept of 'automated decision-making' from the General Data Protection Regulation (GDPR) is relevant.[22]

The above consideration is supported by paragraphs (52)-(59) of the guidelines. The role of human involvement in the creation of generated output by (high-impact) algorithms is discussed in more detail in 5.3. Does the algorithm have a significant effect on the outcome of the process?

The examples (predictions, content, recommendations or decisions) are an important signal what the legislator considers as the output of an AI system during use. See also paragraph (28) of the guidelines. On the basis of this list, a number of types of algorithms that do not qualify as AI systems can be excluded. For example, we find that algorithms that calculate descriptive (population) statistics, such as averages and standard deviations, are not an AI system. When calculating the average income of a group of individuals, the output is not a 'prediction, content, recommendation, or decision'. This is confirmed by paragraph (46) of the guidelines. When a statistical model is used to estimate a score for a new data point, it is a prediction. According to this line of reasoning, simple data processing and visualization systems, such as dashboards that display population statistics, do not count as AI systems. This is confirmed by paragraph (47) of the guidelines.

We therefore see the types of output of an AI system as an important characteristic that distinguishes AI systems from algorithms, specifically in combination with and in relation to the concepts of autonomy and inference. The type of output generated by an algorithm is therefore the first question in the dynamic questionnaire to identify AI systems. See 6.1 Q1 – What type of output does the application derive?

For the question whether an algorithm with a 'prediction, content, recommendation, or decision/decision' as output is actually an AI system, it is important to check how the output is created. The process through which the output is obtained in light of the AI system definition is further analysed in 3. Inference.

[19] Supra note 11

[20] See also advice on automated decision-making, Dutch Data Protection Authority https://www.autoriteitpersoonsgegevens.nl/documenten/advies-geautomatiseerde-besluitvorming

[21] How 'algoprudentie' can contribute to the responsible use of machine learning algorithms, A. Meuwese, J.Parie, A.Voogt, 2024, Nederlands Juristenblad (NJB) https://algorithmaudit.eu/nl/knowledge-platform/knowledge-base/white_paper_algoprudence/

[22] Art. 22 GDPR. See also note 16 above.

2.8 Physical and virtual environment

Recital 12 states that "*can influence physical or virtual environments*" has the following meaning from the AI system definition:

"For the purposes of this Regulation, environments should be understood to be the contexts in which the AI systems operate, whereas outputs generated by the AI system reflect different functions performed by AI systems and include predictions, content, recommendations or decisions." – see recital 12 sentence 10.

The physical and virtual environment are complementary. The combination of the two environments is exhaustive. So, this refers to systems that exert any influence on any environment. This excludes systems that have no influence at all, for example because they have not yet been put into use. This reading is confirmed by paragraph (60) of the guidelines. Furthermore, none of the sources offer a helpful explanation for the concept of 'influence'. There seems to be almost no system imaginable that is employed but does not influence an environment.

In any case, the requirement of 'influence on the physical or virtual environment' is not a characteristic that distinguishes AI systems from algorithms. The concept of influence is further discussed indirectly in the concepts in 3. Inference and 4. Autonomy.

3. Inference

The capability to infer is the most important element of the definition to distinguish AI systems from other algorithms. This section analyses several passages from recital 12 and relates them to the definition of an AI system.

Recital 12 states that the power of inference has the following meaning:

"A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to derive models or algorithms, or both, from inputs or data." – see recital 12 sentence 3-4.

"The techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives, and logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved." – see recital 12, sentence 5.

"The capacity of an AI system to infer transcends basic data processing by enabling learning, reasoning or modelling." – see recital 12, sentence 6.

The first and last sentences frame the interpretation: the inference capacity is an important characteristic by which AI systems can be identified and it is specifically this characteristic that distinguishes AI systems from other data processing by *"learning, reasoning or modelling"*. Note that only one of these three characteristics is required: learning, reasoning or modelling.

On the basis of these three key concepts, the above sentences from recital 12 are analysed.

3.1 Learning and modelling

Recital 12 states that inference capacity relates to:

"a capability of AI systems to derive models or algorithms, or both, from inputs or data." – see recital 12, sentence 4.

When models or algorithms are derived from data, it is modelling or learning. Examples include learning the weights of a neural network used for speech recognition or an algorithm that automatically selects features for profiling. Different experts use different terms for this, such as learning, modelling, training or fitting. See also the examples mentioned about (un)supervised, reinforcement and deep learning in paragraphs (33)-(38) of the guidelines. Regardless of the terminology used, it follows from this passage of recital 12 that inference occurs when a model or algorithm is derived from input or data. From this passage it follows that AI systems must have the ability to derive. By this we mean that there must be a degree of automation in deriving models or algorithms from data, specifically in the development phase of the AI system. See also paragraph (28) of the guidelines. When a data analysis is first performed, for example to determine the average age of a population, which serves as input for domain experts who manually create an algorithm, then there is no situation in which an AI system derives an algorithm from data.

Recital 12 further states:
"The techniques that enable inference in building an AI system include machine learning-based approaches ..." – see recital 12 sentence 6.

In machine learning, a model is 'learned' from a dataset, often called training data. In many cases, statistics are used to calculate model parameters that best fit the available dataset. For data scientists, calculating parameters based on input data is best expressed as the .fit() function, as used in scikit-learn and statsmodels Python software. Calculating an average, using a simple formula, is an example of a parameter. Other examples are calculating linear regression coefficients, using a more elaborate formula, or the weights of a neural network using a very complex formula.

Machine learning also involves learning the variables and thresholds of a decision tree for regression and classification. This could mean learning a simple decision tree, but also learning groups of decision trees, such as ensemble-based tree learning. Such as: random forest, xgboost, explainable boosting etc. These are all examples of machine learning. This is confirmed by paragraphs (30) to (33) of the guidelines.

Whether a data-driven calculation of model parameters is called machine learning differs per domain expertise. An econometrician or statistician would probably not call the development of a linear model, such as a regression analysis or general linear model (GLM), machine learning. However also in these cases, a model is derived from an available dataset. Based on the text of recital 12, we do not see further guidance as to which statistical methods do or do not qualify as machine learning. The specific definition of machine learning appears to be irrelevant to the definition of an AI system. We conclude that all cases when a model is fitted, trained or learned from data fall under the concept of inference.

Just deriving model parameters or rules from input data, for example learning regression coefficients, does not make a model or algorithm an AI system. Recital 12 states that inference capacity refers to:

a) *"the process of obtaining output predictions, content, recommendations or decisions ... which can influence physical and virtual environments."*;
b) *"a capability of AI systems to derive models or algorithms, or both, from inputs or data."* – see recital 12 sentence 4.

When learning regression coefficients, b) but not a) is met. After all, when learning regression coefficients, no predictions are made for new data points. Point a) is about applying the learned model or algorithm to new data. This process is referred to by data scientists as .predict(), as used in scikit-learn and statsmodels Python software. This also relates to the output of an AI system specified by the legislator, namely: "predictions, content, recommendations or decisions". Only after applying this .predict() function output is generated that is required by the definition of an AI system. In the case of recommendations and decisions, a score is often first predicted, using a learned model, after which a recommendation or decision is made on the basis of this score. A model – based on statistics or machine learning – is an AI system when model parameters or rules are derived from data and then a prediction or similar follows. See 2.7 Predictions, content, recommendations or decisions.

'Generating output' is an important factor in distinguishing AI systems from algorithms and is therefore included in the dynamic questionnaire. The first question from the questionnaire is related to the generated output. The second question is about deriving models or algorithms from data. Through a follow-up question, in which the user must provide an explanation which (statistical) methods are used, it can be assessed on a case-by-case basis whether an AI system is involved. See 6. Dynamic questionnaire.

3.1.1 Contradictions in the European Commission's guidelines

As explained in the previous sections, paragraphs (1)-(41) of the guidelines support our view on the definition of an AI system. This is not the case for paragraphs (42)-(45) and (49)-(51). These paragraphs introduce exceptions that are difficult to reconcile with our above analysis learning and modelling. The exceptions from the guidelines are motivated by arguments that contradict the legislative text of the Act. This subsection explains which specific passages introduce ambiguity.

Paragraphs (42)-(45) from section 5 of the guidelines argue that: "*Systems used to improve mathematical optimisation or to accelerate and approximate traditional, well established optimisation methods, such as linear or logistic regression methods, fall outside the scope of the AI system definition. This is because, while those models have the capacity to infer, they do not transcend 'basic data processing*". The first sentence of this exception focusses on the application of an algorithmic system– namely to improve optimization – rather than on properties of the system itself. However, the definition is only about characteristics of the system. The way in which a system is used is not part of the definition and can therefore not be used as an argument for an exception. The second sentence is even more directly contradicts the legal text. It follows from recital 12 of the AI Act that "*the capacity of an AI system to infer transcends basic data processing* … ". Paragraph 42 of the guidelines therefore contradicts the description of the inference capacity as explained in recital 12 of the AI Act.

Confusion also follows from the argument cited in these paragraphs that: "*an indication that the system does not exceed the basic processing of data can be deduced from the fact that the method has been used for many years*". Statistically, it does not matter whether the method used by an AI system has been used for a long or short time. As explained earlier, linear and logistic regressions are elementary forms of machine learning that can be used to generate predictions. Neural networks are also decades old technology. It seems untenable to place these systems outside the scope of the AI Act on the basis of the 'established methods' argument.

In paragraph (48) it is suggested that a lack of adaptability is a reason for exempting systems from the definition. This contradicts paragraphs (42) to (43) of the same guidelines, which explicitly state that adaptability is not a prerequisite for the definition of an AI-system. The lack of an optional feature cannot be used to argue that a system is not an AI system.

Paragraphs (49) to (51) state that all machine-based systems that function on the basis of simple rules – that could have been established through statistical learning and thus involve machine learning – do not qualify as AI systems, 'because of their performance'. This passage also contradicts the legislative text. In the definition of AI system and recital 12, complexity of the system or performance is not attributed as a characteristic of inference. However, sentence 5 of recital 12 explains that *"the techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives"*. Machine learning methods that result in simple rules, such as linear regression or decision tree learning, are also used to learn how to achieve an objective on the basis of data. Putting forward arguments about complexity and performance are not only new with regard to the legislative text to exempt systems from the AI Act, they also bring additional ambiguity: when is a system 'simple' enough to fall under this exception?

As an example of simple systems that do not qualify as AI systems, paragraphs (50)-(51) cite an 'average baseline prediction' and a 'static estimation system'. However, these methods are not a form of machine learning – no model parameters or decision rules are derived from data. According to the definition in the AI Act, these examples would not qualify as an AI system because there is no learning or reasoning. The reference to machine learning in paragraph (49) creates unnecessary confusion about the meaning of machine learning and how it relates to the definition of an AI system.

In view of paragraph (7) of the guidelines – from which it follows that the guidelines have a lower legal status than the legal text – this white paper follows the legislative text of the AI Act to determine the scope of an AI system. The above exceptions have not been included in our interpretation of the definition due to their inconsistency.

3.2 Reasoning: logic and knowledge-based approaches

Inference power can also refer to the ability of an AI system to reason – see recital 12 sentence 6. This shows that there is a type of systems in which there is no learning or modelling, but there is inference.

This raises the question: what type of algorithms involve reasoning? Recital 12 mentions several examples of systems that are not covered by this: *'rules established exclusively by natural persons for the purpose of performing actions automatically'* and *'basic processing of data'* – see recital 12 sentence 2 and 6.

Recital 12 offers little additional explanation for the concept of 'reasoning'. Recital 12 does mention the following:

"The techniques that enable inference while building an AI system include... logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved." – see recital 12 sentence 5.

Logical and knowledge-based approaches to AI do not involve machine learning, they involve inference because they involve reasoning.

Logic- and knowledge-based approaches to AI are also referred to as symbolic AI in academia and in the OECD memorandum.[23] Symbolic AI has been used since the 80s and 90s, for example, in chess computers or medical decision support systems. However, with the great advancements in machine learning, deep learning, and generative AI, less and less attention has been paid to this form of AI.

Recital 12 does not provide additional information on the definition and interpretation of logic- and knowledge-based approaches to AI systems. The original proposal of the AI Act does include additional clarification: *"Logic- and knowledge-based*

[23] Supra note 11

approaches, including knowledge representation, inductive (logical) programming, knowledge bases, inference and inference machines, (symbolic) reasoning and expert systems".[24] These examples are in line with interpretations of symbolic AI in academia. These examples are also included in the guidelines in paragraph (39).

To distinguish between logic- and knowledge-based AI systems from algorithms, we need to distinguish what makes these techniques different from *"rules defined solely by natural persons to automatically execute operations"* and *"the basic processing of data"*. We explain logic- and knowledge-based approaches further on the basis of two academic standard works in AI: Artificial Intelligence by Russel and Norvig and Artificial Intelligence by Poole and Mackworth.[25] In summary, logic- and knowledge-based approaches to AI consist of a knowledge-base and reasoning component. Paragraph (30) of the guidelines confirm this reading.

i) **Knowledge-base:** An explicit representation of (domain) knowledge. This knowledge consists of, for example, rules, facts and relationships. See paragraph (39) of the guidelines. Logic is often used for this, in which knowledge is expressed in propositions and connectives, such as ¬A, A∧B, A∨B, where a proposition (e.g., A) can only be true or false. Other well-known forms of knowledge bases are knowledge graphs.
ii) **Reasoning component:** This component defines how the system can reason about the knowledge in the knowledge base and input data, for example by means of formal logic. This component is also called an inference engine, for example a deductive or inductive engine. Reasoning can also take place through operations such as sorting, searching or matching. See also paragraph (39) of the guidelines. By means of the reasoning component, new knowledge and new rules can be derived.

The guidelines do not explicitly mention this distinction between a knowledge base and a reasoning component. It follows from paragraph (39) that: *"AI systems learn from knowledge including rules, facts and relationships encoded by human experts. Based on the human experts encoded knowledge, these systems can 'reason' via deductive or inductive engines or using operations such as sorting, searching, matching, chaining."*

Both components of logic- and knowledge-based are carefully built up and require a lot of domain knowledge. These approaches are often used when there is a large amount of fixed knowledge and rules in a domain, which can then be reasoned on. Think of a medical decision support system where the knowledge base contains medical facts about symptoms, diagnoses and possible treatments, the reasoning system can then propose a possible treatment based on input data of symptoms.

Strictly logic- and knowledge-based approaches to AI are a rarely applied. Nowadays, these techniques are commonly used in combination with machine learning. In that case, the system would be an AI system due to the use of machine learning, see 3.1 Learning and modelling. Developers who use this type of technology are probably aware that they are using this type of AI system. We see the 'logic and knowledge-based approaches' as an important characteristic to distinguish AI systems from algorithms, only in those rare cases where ML is not used.

[24] See Annex I of Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=CELEX:52021PC0206

[25] Artificial Intelligence: foundations of computational agents. Poole, D.L. and Mackworth, A.K., 2010. Cambridge University Press. Artificial intelligence: a modern approach. Russell, Stuart J., and Peter Norvig. Pearson, 2016. For an understandable explanation, see also: https://en.wikipedia.org/wiki/Knowledge-based_systems

Logic- and knowledge-based approaches are a factor that distinguishes AI systems from algorithms. A question about it has therefore been included in the dynamic questionnaire. See 6.3 Q3 – Is the application automation of human-defined rules?

3.2.1 Reasoning, coded knowledge and rule-based systems

There are no other approaches referred to in the AI Act that relate to reasoning, other than logic and knowledge-based approaches.

It can be argued that in the case of a simple manually defined rule-based algorithm, there is reasoning. However, this is incompatible with the objective of the definition: *"the definition should be based on the key characteristics of AI systems that distinguish it from simpler traditional software systems or programming approaches..."*. When rule-based algorithms reason, all types of software systems reason and that goes against the previous sentence. Regardless of whether there is reasoning, *"rules defined solely by natural persons"* are not an AI system – recital 12 sentence 2.

The passage on 'encoded knowledge' – recital 12 sentence 6 – should therefore be seen in the light of logic- and knowledge-based approaches. In this context, encoded knowledge relates to the form in which knowledge is encoded in a knowledge base, as described above. Rule-based algorithms, in which human knowledge is encoded, are not applied in practice by means of a knowledge base (also known as a 'knowledge-based approach'). The passage *"coded knowledge*" therefore does not refer to rule-based algorithms that we know from practice.

4. Autonomy

Recital 12 states that '*different levels of autonomy*' from the AI system definition has the following meaning:
"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." – see recital 12 sentence 11.

In order to meet the 'autonomy' requirement, there must be some degree of autonomy, as also discussed in the 2.3 Varying levels of autonomy and paragraph (14) of the guidelines.

'A certain degree' is a weak requirement: a system does not have to be completely autonomous to meet this requirement.

The OECD memorandum states that: "the autonomy of an AI system refers to the extent to which a system can learn or act without human involvement". This implies that every learning algorithm is autonomous to a certain extent. In other words, if the inference requirement is met, the autonomy requirement is also met. Furthermore, the OECD memorandum links autonomy to the different types of outputs generated, with decisions being the most autonomous and forecasts the least autonomous. From this formulation we conclude that the OECD also considers predictions to be autonomous to 'some extent'.

The guidelines state in paragraph (18) that a system *"that requires manually provided inputs to generate an output by itself is a system with 'some degree of independence of action'"*, i.e., a certain degree of autonomy. This lower limit applies to all systems that generate the output types from input data mentioned in earlier sections. With a consideration of the type of output of an algorithm (2.7 Predictions, content, recommendations or decisions) and the inference capacity (3. Inference) can therefore also be used to meet the autonomy requirement.

All in all, we conclude that the 'autonomy' requirement has no additional capacity compared to the aforementioned other characteristics, to distinguish AI systems from algorithms. However, human intervention does play an important role in distinguishing high-impact algorithms from regular algorithms. A question about human intervention is therefore included as a question in the dynamic questionnaire. See 6.7 Q7 – Which of the following options best describes the effect of your application on the outcome of the application?

Paragraph (21) of the guidelines emphasises that a lack of human intervention entails additional risks and therefore requires additional measures. This is in line with the concept of 'automated decision-making' as used in the General Data Processing Regulation (GDPR) and recent advice published on this subject by the Dutch Data Protection Authority (DPA).[26]

5. High-impact algorithms

Algorithms[27] is a term that has been used for some time by the Dutch government to refer to a broader category of automated systems, which also includes AI. Algorithms can have significant impact on data subjects, even when they are not AI systems. An example is profiling algorithms whose decision rules have been defined solely by human experts. An example is the algorithm of the Dutch Executive Agency for Education (DUO) that was used in the period 2012-2023 to assign students a risk score for violation of the requirements for a college grant to cover living expenses.[28] See 2.1 Interpretation of the definition of an AI system using recital 12. These algorithms with an increased risk of impacting data subjects are referred to as high-impact algorithms.

High-impact algorithms appear to be more prevalent than AI systems, particularly in the public sector. In a November 2024 report by the Dutch Ministry of Finance, it was stated that the Dutch Tax Authority, Customs Agency, Benefits agencies and the Ministry itself all do not have any high-risk AI systems in use.[29] In contrast, these organizations reported the use 77, 55, 22, and 8 high-impact algorithms, respectively. 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 to them. This white paper and accompanying dynamic questionnaire therefore focus not only on identifying AI systems, as defined in the AI Act, but also on identifying high-impact algorithms.

Whereas the previous sections focus on the question of how to identify AI systems, this section focuses on the question of how to distinguish 'high-impact algorithms' from other algorithms. An explanation follows how high-impact algorithms can be recognized. For this goal the Guideline Algorithm Register[30] of the Ministry of the Interior is followed. High-impact algorithms must be published in the Dutch national Algorithm Register, unless there is a ground for exception.[31]

26 Art. 22 GDPR; Advice on automated decision-making, Dutch Data Protection Authority (2024); State Attorney's Advice on Automated Selection Techniques, Pels Rijcken.

27 'Algorithm' as defined by the Dutch Court of Audit (2021): "A set of rules and instructions that a computer automatically follows when making calculations to solve a problem or answer a question".

28 Preventing prejudice, Algorithm Audit (2024)

29 Parliamentary Papers II 2024-25 2025D00512

30 Guideline Algorithm Register of the Ministry of the Interior and Kingdom Relations (2023).

31 Grounds for exception as stated in the Guideline are: "legal grounds for exception as specified in the Open Government Act (Woo) and the Public Health Act (Wpg), or 'gaming the system'."

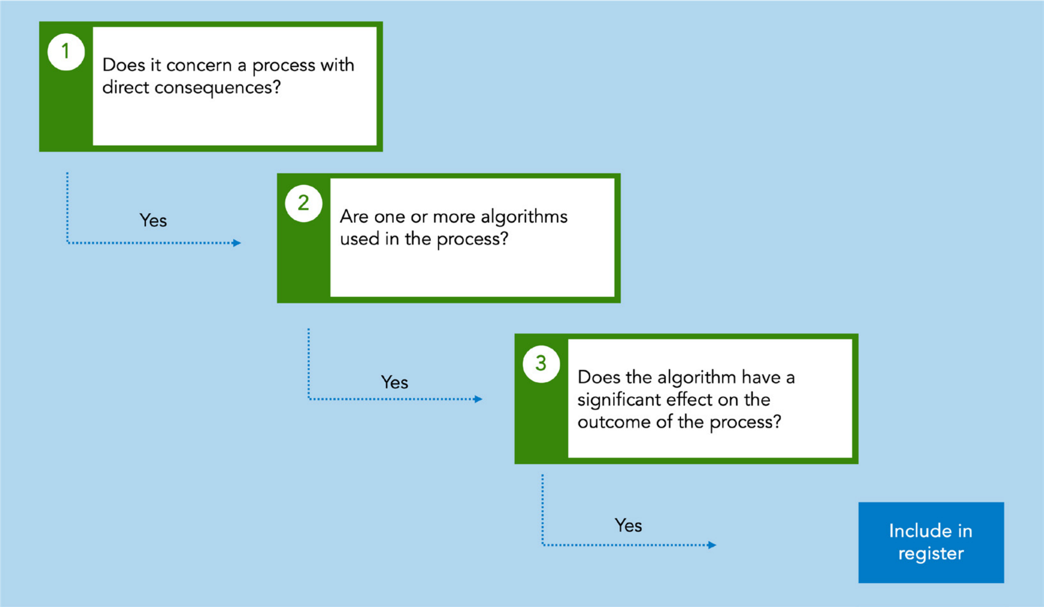


Figure 2 - Questions from the Guideline Algorithm Register that can be used to determine whether an algorithm is a 'high-impact algorithm'.

Based on this analysis, the dynamic questionnaire for the identification of AI systems is supplemented with four questions on the basis of which high-impact algorithms can be identified. See section 6.4-6.7 in 6. Dynamic questionnaire.

The definition of a 'high-impact algorithm' from the Guideline Algorithm Register is as follows:
- **Direct consequences:** the algorithm has direct consequences for those involved (citizen, organization), such as: imposing a fine or refusing a subsidy; or
- **Classification:** the algorithm influences how the government categorizes or approaches a data subject or group, such as: profiling or risk indication for control.

The above categories are explained in the Guideline Algorithm Register on the basis of three questions. See Figure 2.

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?

Note that the Guideline first examines the impact of the process on those involved (citizens, organisations) before asking about the effect of the algorithm on that process.

5.1. Is it a process with direct consequences?

A high-impact algorithm is used in a process that has direct consequences for those involved.

1. *"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."*
2. *"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."*
3. *"Stakeholders include everyone who has to deal with the Dutch government. We summarize this as citizens and organizations."*

Direct consequences are broadly defined, making many algorithms potentially high-impact algorithms. To determine whether there are direct consequences in the algorithm-driven decision-making process, it is first checked whether there is a decision. A decision must be interpreted broadly. Not only a formal decision, as defined in the Dutch Public Administration Law (Awb art.1:3), has an impact on citizens and organisations, other decisions can also have significant consequences for those involved and can therefore belong to the category of high-impact algorithms. That is why decisions are discussed in the remainder of this paper and in the dynamic questionnaire.

This can either involve decisions for individual stakeholders, or the approach or categorisation of groups of stakeholders. Once it has been established that there is a decision-making process for individual citizens, it must be examined what kind of decision is taken in the algorithm-driven decision-making process (prioritisation, a decision on a formal complaint or objection, a decision with financial consequences, etc.). In this way, it is determined whether this decision has a direct impact. For these aspects, three questions have been included in the questionnaire. See 6.4 Q4 – Is in the process a decision made for individual citizens or civil servants?, 6.5 Q5 – What type of decision is made in this process? and 6.6 Q6 – Does the process contribute to how the governmental institution categorizes or approaches (groups of) citizens or civil servants?

5.2. Are one or more algorithms used in the process?

A high-impact algorithm refers to a process in which one or more algorithms are used.

1. *"One or more algorithms are used in the process.";*
2. *"A process often consists of several steps, some of which are carried out by algorithms and some by humans.";*
3. *"An organization often knows best which set of steps together form the process with impact on those involved.";*

It is assumed for the purpose of this analyses that at least one algorithm is involved in the decision-making process. It is not a distinguishing factor to separate high-impact algorithms from other algorithms and is therefore not included as a question in the questionnaire.

5.3. Does the algorithm have a significant effect on the outcome of the process?

A high-impact algorithm has a significant effect on the outcome of the process.

1. *"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."*
2. *"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 process. The weighting factors are filled in by the space or freedom that an administrative body is entitled to in carrying out its tasks."*

A significant effect on the outcome of the process can be read as: how directly is the outcome of the process connected to the output of the algorithm? There is also another implicit element in this: would the process be different without the algorithm? There is a parallel here with the definition of AI system in the AI Act. In it, autonomy is linked to the different types of outcomes and the degree of human intervention.

The Guideline explicitly states that an *"algorithm [that] automates/digitizes a manual work instruction"* is not a high-impact algorithm. These manual instructions can be divided into a) laws and regulations and formal policy, and b) more informal work instructions including other human-defined rules. Category a) is formally laid down in laws and regulations and is therefore subject to democratic and institutional control. The automation of these rules does not change the outcome of the process. This type of algorithms is therefore less impactful than algorithms for which the rules are not so clearly specified.

However, laws and regulations or formal policy do not always contain provisions that can be implemented directly in a decision rule. Sometimes there is '0.8-to-1 automation', where there is room for the organization to interpret a provision itself and formulate it as a decision rule.[32] In that case, this does not qualify as a AI system since "it concerns rules that have been adopted exclusively by natural persons for the purpose of performing actions automatically" (recital 12 AI Act) and therefore falls outside the scope of the AI Act. This type of algorithms belongs to the category b) described above.

An example of this type of algorithm is the risk profile used by DUO in the CUB-process. This algorithm consisted of simple rules defined solely by natural persons. These rules could also have been executed manually as a work instruction. Yet, this algorithm led to indirect discrimination and ultimately the compensation of more than 10,000 students.[33] Although this category does not qualify as a high-impact algorithm according to the Guideline, this algorithm appears to be very impactful. That is why this paper deviates from the Guideline on this point. When an algorithmic application is an automation of rules that are not referred to in policy or regulation, we see this application as a potentially high-impact algorithm.

The effect of an algorithm on the outcome of a decision-making process is an important factor to distinguish high-impact algorithms from other algorithms and is therefore included as a question in the dynamic questionnaire. This involves asking

[32] An example from the Unemployment Insurance Act (WW) can be found in 'The A stands for algorithm: how to strengthen the Awb in this area' O.A. al Khatib, M.H.A.F. Lokin, R.J.H. Bruggeman & A.C.M. Meuwese (2024).

[33] Supra note 6

whether there is (meaningful) human intervention and in what form. See 6.7 Q7 – Which of the following options best describes the effect of your application on the outcome of the application? In addition, the type of outcome also determines the effect of the algorithm on the outcome of the process. See 6.1 Q1 – What type of output does the application derive? The type of algorithm described in a) that is a direct automation of a manual work instruction is incorporated in the 6.3 Q3 – Is the application automation of human-defined rules?

6. Dynamic questionnaire

The above analyses of the definition of an AI system and of a high-impact algorithm have resulted in a dynamic questionnaire to identify these types of algorithms. Within three questions it can be determined whether an application is an AI system. These questions are introduced and explained in 6.1-6.3. By asking at maximum four additional questions, it can be determined whether an application is a high-impact algorithm. These questions are introduced and explained in 6.4-6.7.

The questionnaire has four outcomes:

> **Not in scope:** algorithm is not an AI system, nor is it a high-impact algorithm;
> **High-impact algorithm:** algorithm is a high-impact algorithm, but not an AI system;
> **AI system:** algorithm is an AI system, but not a high-impact algorithm;
> **High-impact algorithm and AI system:** algorithm is both an AI system and a high-impact algorithm.

An overview of all the results can be found in Figure 3. A flowchart of all questions and outcomes can be found in Appendix B.

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

As explained in 2.7 Predictions, content, recommendations or decisions the output generated by an algorithm contains important information about whether the algorithmic application may involve an AI system. Because algorithm developers, product owners, line managers, and other executive users

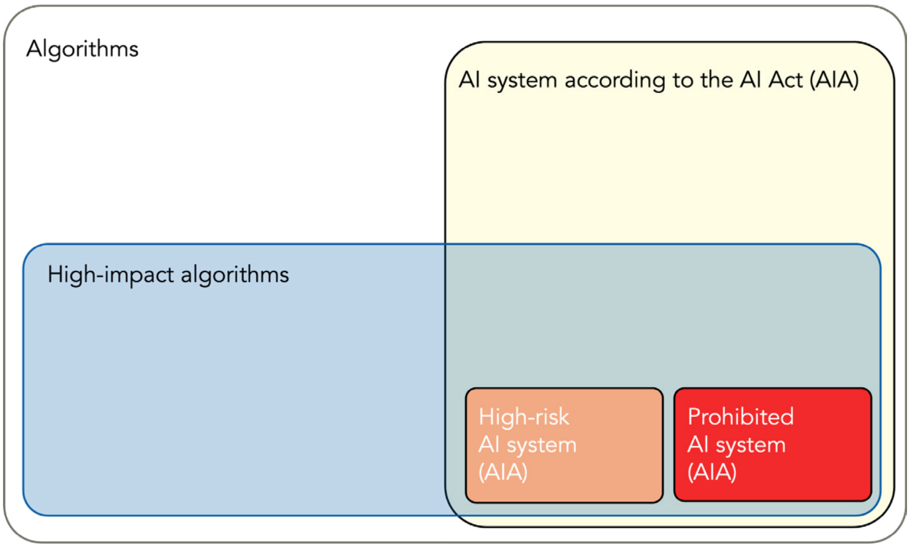


Figure 3 - Venn diagram of all outcomes of the dynamic questionnaire

are typically familiar with the output of algorithms, and they are the target audience that uses the tool, the dynamic questionnaire starts with a query about the output that algorithmic application produces. See Figure 4.

The AI Act mentions "predictions" as a possible form of output. Prediction is a broad concept that is not interpreted in the same way by everyone. In data science, a prediction does not have to be about the future. A prediction can also relate to a data point that has not been observed before. The answer options include explanatory terms for types of predictions that may be recognizable to users of the questionnaire (score, ranking, label, object-, face- or voice recognition). Despite the fact that every score, ranking, label, classification, and image recognition is essentially a prediction, it favours accessible language over the prevention of duplication.

Because dashboards are a common data-driven application that raises questions about the scope of the AI Act, it has been included as a separate answer option. Typically, dashboards are only used for data visualization and therefore do not meet the definition of an AI system. These dashboards do not transcend *"the elementary processing of data by enabling learning, reasoning or modelling."* – see recital 12 sentence 6 and 3. Inference. Underlying algorithms, the outcome of which is shown on a dashboard, can be an AI system, but this does not apply to the dashboard itself. If a dashboard visualizes scores, rankings or similar, a different answer must be filled in at Q1 with which the applications may fall within the scope of the AI Act. Based on the 'dashboard' answer options, it is therefore concluded that there is no question of an AI system.

The answer options – score, ranking, label, recommendation, decision, content, object-, face-or voice recognition – are also helpful in distinguishing algorithms from high-impact algorithms. Algorithms with this type of output have an effect on the outcome of the process and are therefore a high-impact algorithm. These forms of output are in line with the description given in the AI Act: *"predictions, content, recommendations or decisions"*.

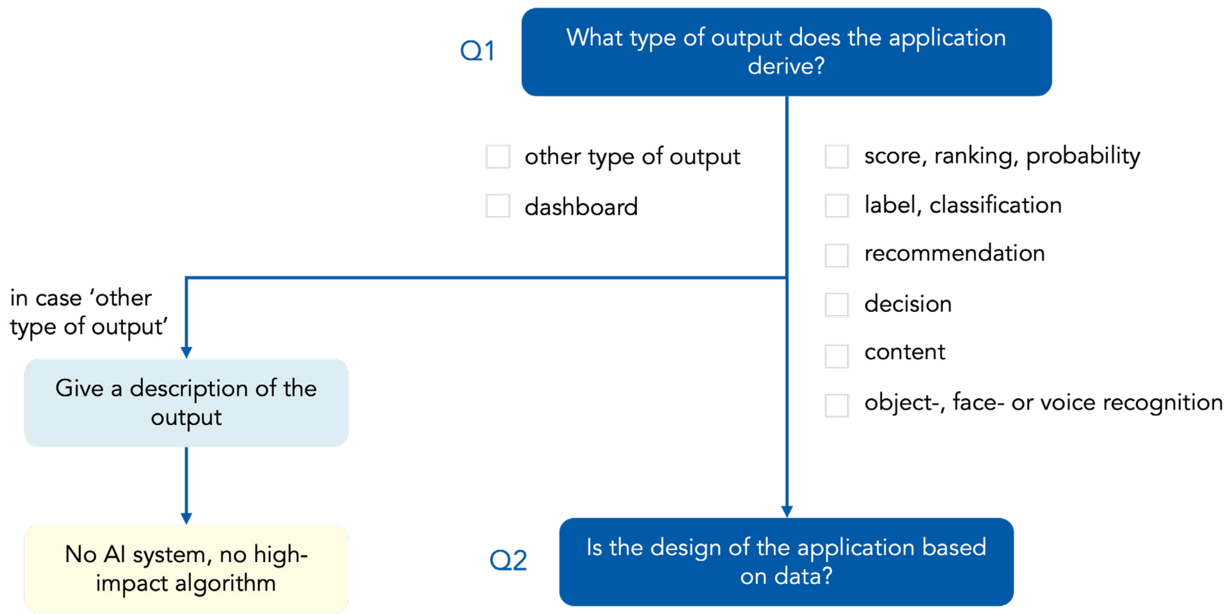


Figure 4 - Q1 asks what type of output the algorithmic application generates.

There is also no question of an AI system or high-impact algorithm when the answer option *"other type of output"* is chosen. In this case, the user is asked to provide a description of the output. All other answer options may involve an AI system and a high-impact algorithm. In this case, the user will be redirected to Q2.

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

Where Q1 queries the output of the algorithmic application, Q2 focuses on the way in which the output is generated. As in 3. Inference *"the ability of AI systems to transmit models or algorithms, or both, to be derived from input or data"* an important factor in distinguishing an algorithm from an AI system. After the user selects one of the answer options that lead to Q2 in Q1, it is asked whether the design of the algorithmic application is based on data.

If the application contains components derived from data, then the application 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. See 3.1 Learning and modelling.

Based on the guidelines, it follows that not all applications whose components are derived from data are an AI system. Exceptions to this rule are discussed in 3. Inference. To make this assessment, the answer option 'yes, the application contains components derived from data' asks the user for an explanation. This information helps to determine on a case-by-case basis whether or not an algorithm is an AI system.

In view of recital 12 of the AI Act, where choices in the design of the application are made manually, the application is unlikely to be an AI system but may be a high-impact algorithm. See 2.1 Interpretation of the definition of an AI system using recital 12. To check whether the application is a high-impact algorithm, the user is referred to Q4.

Even if the design of the application is not based on data, the application can still be an AI system. See 3.2 Reasoning: logic and knowledge-based approaches. To check if this is the case, users are redirected to Q3.

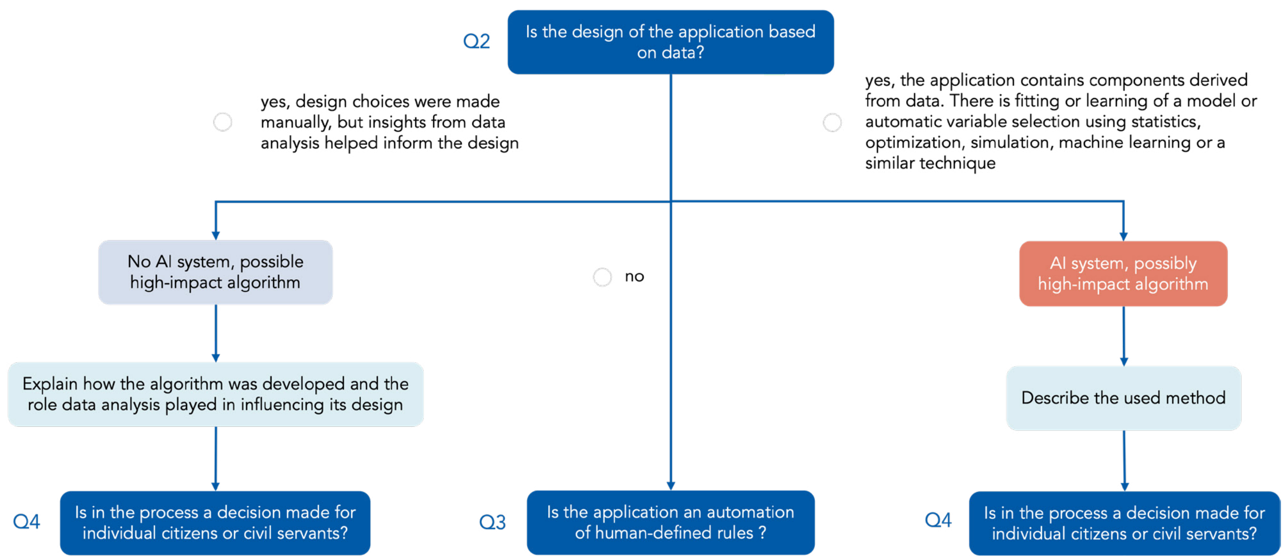


Figure 5 - Q2 is about whether the design of the application is based on data.

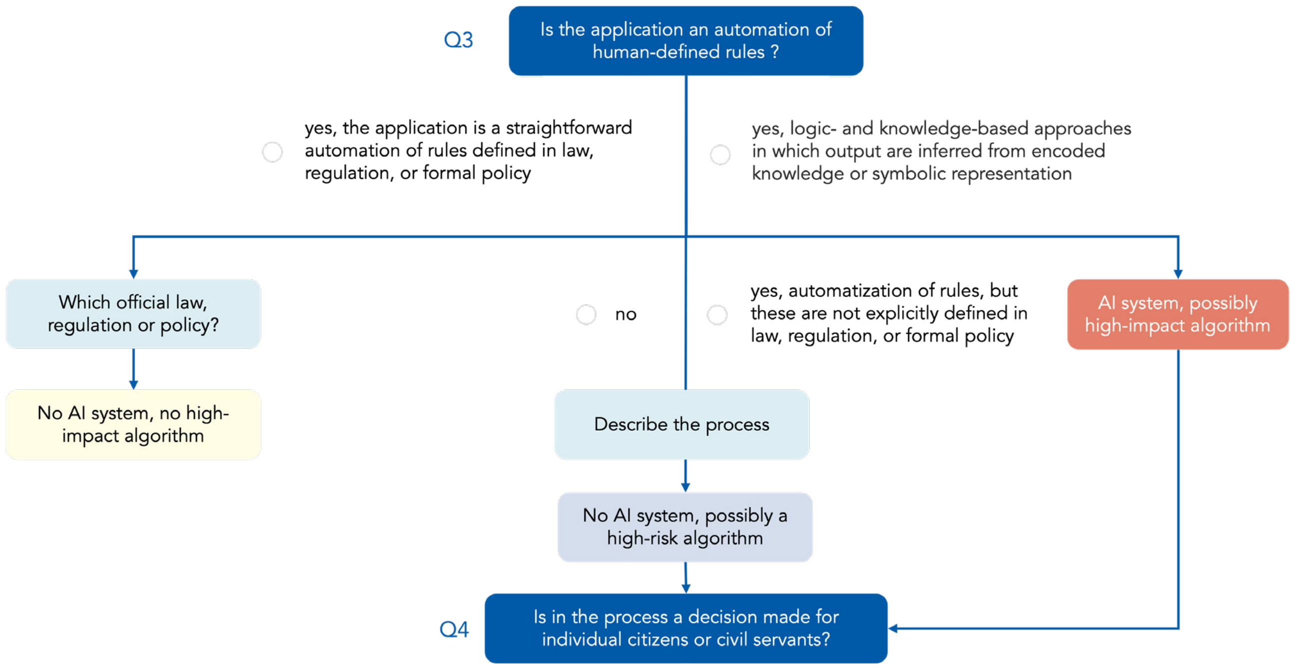


Figure 6 - Q3 examines the extent to which there is human involvement in the creation of rules used in algorithmic application.

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

If rules are used in algorithmic applications, there might not be an AI system involved. See 2.1 Interpretation of the definition of an AI system using recital 12. This will be clarified in Q3.

When rules are created using logic- and knowledge-based approaches, the algorithm does qualify as an AI system. However, logic- and knowledge-based approaches are rare forms of AI. See 3.2 Reasoning: logic and knowledge-based approaches. In other cases where the application is based solely on *"rules defined solely by natural persons"*, then the application is not an AI system. See 2.1 Interpretation of the definition of an AI system using recital 12. If there had been data analysis that played a role in the creation of the decision rule, an earlier question would have had to give a different answer, as a result of which the user would have ended up with Q4 instead of Q3.

Q3 is also particularly important for the identification of high-impact algorithms. The way in which rules defined by natural persons are created partly determines whether an algorithmic application is a high-impact algorithm. We distinguish here between a) laws and regulations and formally established policy, and b) more informal work instructions and other rules drawn up by people. See also 5.3. Does the algorithm have a significant effect on the outcome of the process?

However, laws and regulations or formal policy do not always contain provisions that can be implemented directly in a decision rule. Sometimes there is '0.8-to-1 automation', where there is room for the organization to interpret a provision itself and formulate it as a decision rule.[34] These cases are not one-to-one automation of laws or regulations or formal policies. These systems may be a high-impact algorithm

[34] An example from the Unemployment Insurance Act (WW) can be found in 'The A stands for algorithm: how to strengthen the Awb in this area' O.A. al Khatib, M.H.A.F. Lokin, R.J.H. Bruggeman & A.C.M. Meuwese (2024).

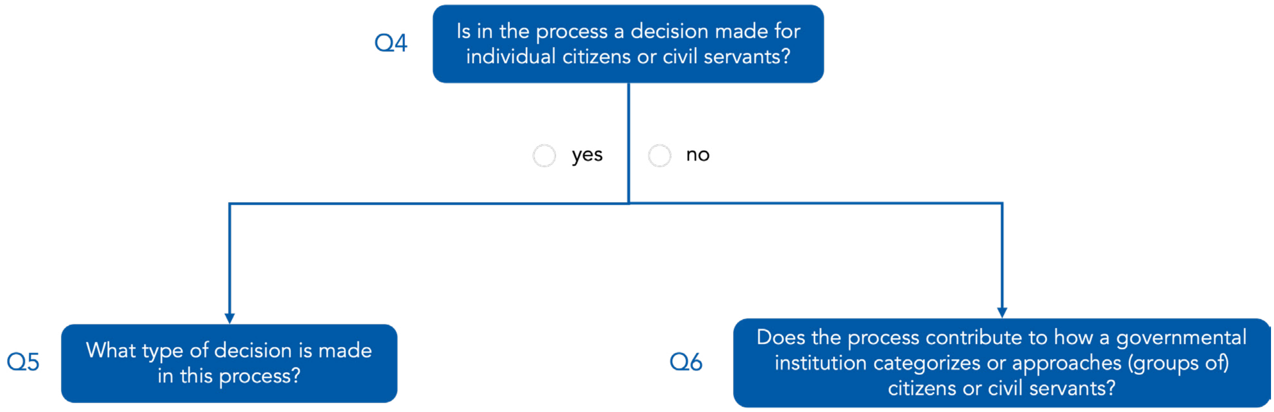


Figure 7 - Q4 examines whether a decision is made in the process in which an algorithm is potentially involved.

When Q3 is answered with 'no', the user is asked to provide a description of the application and is then redirected to Q4.

6.4 Q4 – Is in the process a decision made for individual citizens or civil servants?

On the basis of Q1-Q3, it was determined whether the algorithmic application concerns an AI system. On the basis of a maximum of four additional questions, it can be determined whether the application is a high-impact algorithm. Following the Guideline Algorithm Register, the first question is asked about the process, regardless of the use of an algorithm (Q4-Q6) and only then about the role of the algorithm in this process.

The first question examines whether a decision is taken in the process that relates to an individual citizen, organization or civil servant. This aspect is important to distinguish high-impact algorithms from other algorithms. See 5.1. Is it a process with direct consequences? The emphasis in this question is on individuals. In Q6, the emphasis is on groups.

If a decision is made in the process, it is important to check what kind of decision is made in the process. The user is redirected to Q5.

If no decision is made in the process, the question is asked to what extent the process contributes to the way in which the government categorises or approaches groups of citizens or civil servants. The user is redirected to Q6.

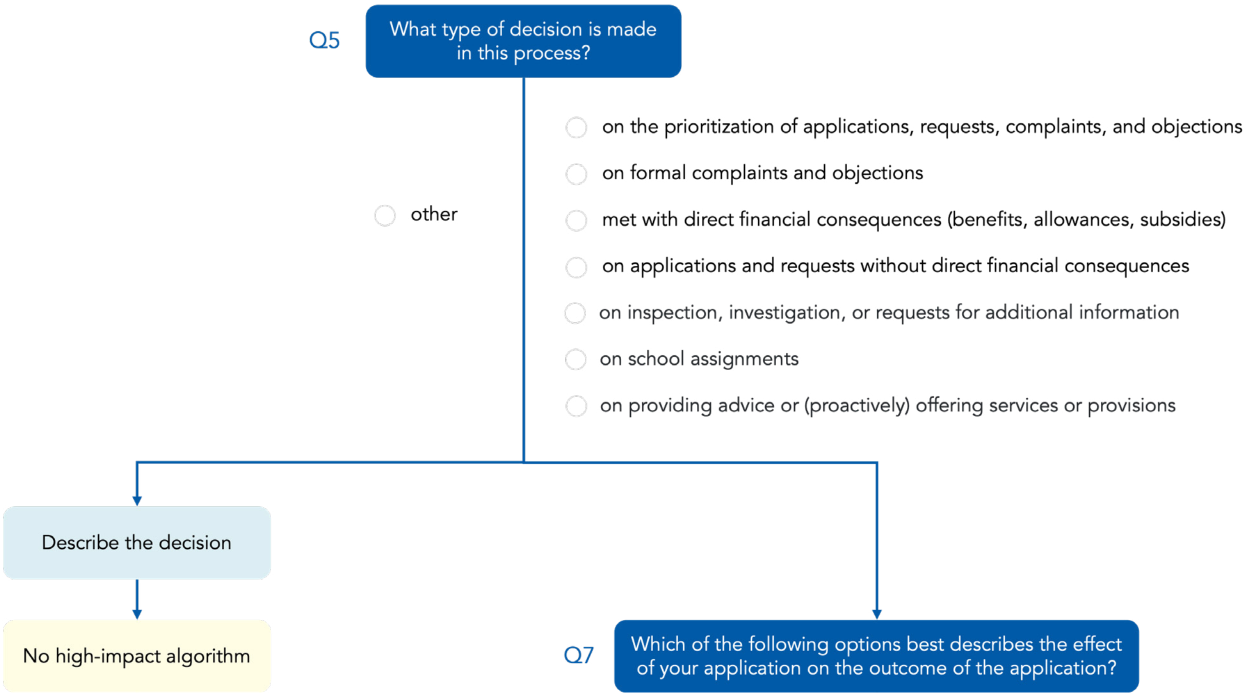


Figure 8 - Q5 asks what kind of decision is made in the process.

6.5 Q5 – What type of decision is made in this process?

Given that a decision in the process is made for an individual citizen, organization or employee, Q5 asks what kind of decisions are made in the process. The type of decision is an important factor in distinguishing high-impact algorithms from regular algorithms. See 5.1. Is it a process with direct consequences?

A decision is broadly defined. This may include decisions on prioritisation of applications, requests, complaints or objections; decisions on formal complaints and objections; decisions with direct financial consequences, such as decisions about benefits, allowances, subsidies, fines, repayments or the possibility of a payment arrangement; decisions on applications and requests without direct financial consequences, such as granting an application for services or granting a licence; decisions on monitoring, research or request for additional information; decision on the allocation of schools; decision on advice to be given or the (proactive) provision of services or facilities. In all these cases, users are redirected to Q7 to determine the effect of the application on the outcome.

When a type of decision does not belong to one of the previous categories, it can be concluded that the impact of the process is limited and that the algorithm in question is not a high-impact algorithm. In this case, a description of the type of decision is requested.

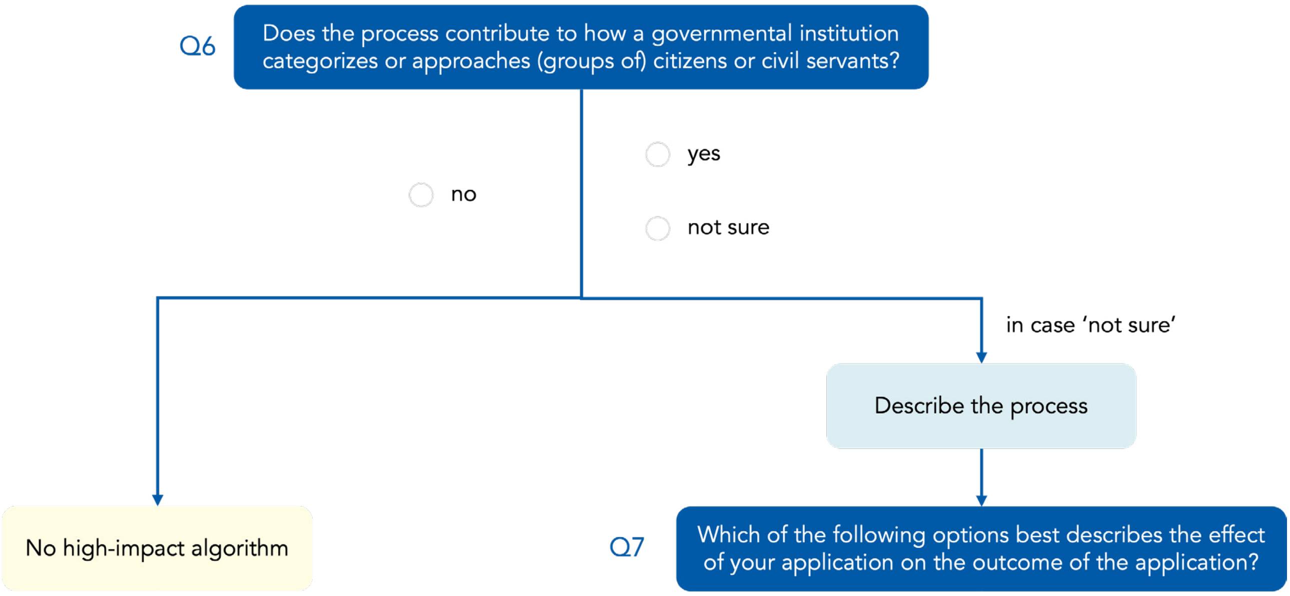


Figure 9 - Q6 examines whether the process contributes to how groups of citizens, organisations or civil servants are categorised or approached by the government

6.6 Q6 – Does the process contribute to how the governmental institution 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, Q6 examines whether the process contributes to how the government categorises or approaches groups of citizens, organisations or civil servants. The importance of this is explained in 5.1. Is it a process with direct consequences?

If the process does not contribute to how the government categorises or approaches groups of citizens, organisations or civil servants, it is concluded that the algorithm in question is not a high-impact algorithm. If this cannot be said with certainty, an explanation is requested, after which the user is redirected to Q7. The user will also be redirected to Q7 if Q6 is answered with 'yes'.

6.7 Q7 – Which of the following options best describes the effect of your application on the outcome of the application?

In all cases where a decision is taken for citizens, organisations or civil servants (individual or groups), it is relevant to find out what the effect of algorithmic application is on the decision-making process. This effect determines whether or not the algorithm is a high-impact algorithm. See 5.3. Does the algorithm have a significant effect on the outcome of the process?

If the outcome of the process is directly determined by the algorithm, without human intervention, then it qualified as a high-impact algorithm. This is also the case when after the completion of the process, human analysts can review the results.

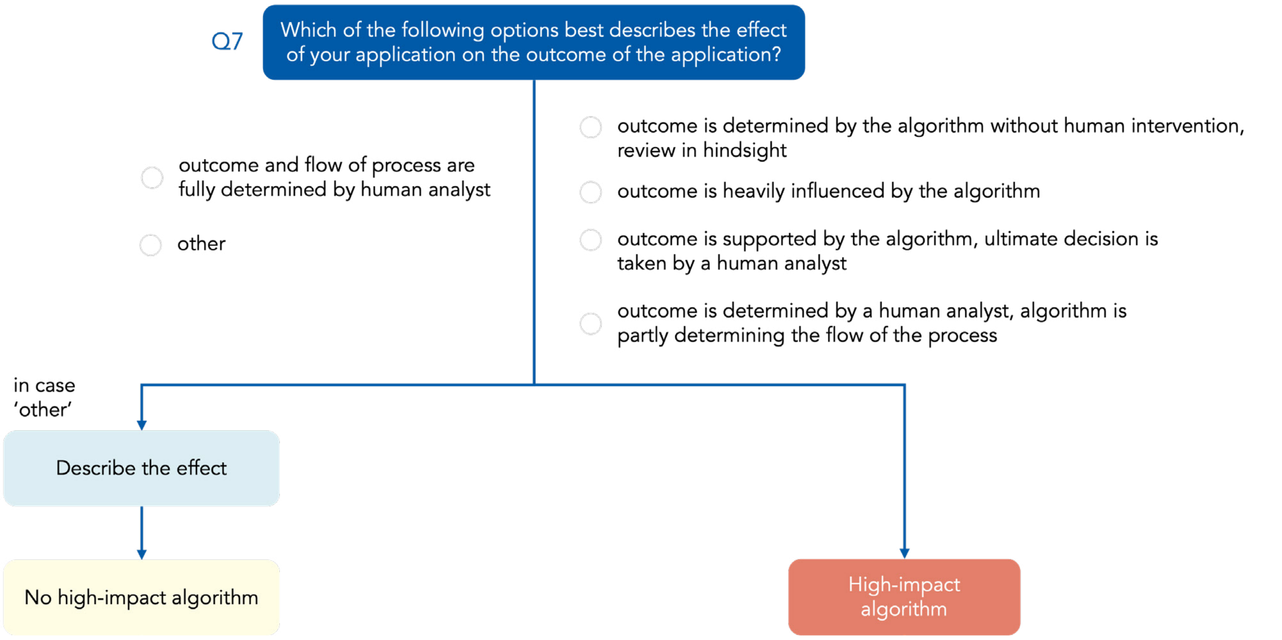


Figure 10 - Q6 examines whether the process contributes to how groups of citizens, organisations or civil servants are categorised or approached by the government

It also concerns a high-impact algorithm when the process is strongly influenced by the algorithmic application. For example, because work instructions determine the consequences of a certain outcome of the application. In this case, a human analyst may make different choices in some cases, but usually the result of the system determines what the end result of the process will be.

Even when the outcome of the process is partly influenced by the algorithm, this qualified as a high-impact algorithm. The result of the application is important for the end result, but the final decision is made by an employee. This employee has the right information, experience/skills, mandate and available time to make the decision.

When the algorithm (partly) determines the course of the process, but the outcome of the process is entirely determined by a human analyst, this also qualifies as a high-impact algorithm. This is the case, for example, when the outcome of the application is a risk score that initiates a control process or a more intensive case evaluation, but the control or evaluation is then carried out entirely by a human analyst.

If the outcome of the process and the course of the process are entirely determined by a human analyst, this does not qualify as a high-impact algorithm. This is also the case when there is a different effect of the application on the outcome. In the latter case, a description of this effect is requested.

Appendix A – Recital 12

Recital 12 of the preamble to the AI Act.

Sentence 1 – analysed in 1. Introduction

The notion of 'AI system' in this Regulation should be clearly defined and should be closely aligned with the work of international organisations working on AI to ensure legal certainty, facilitate international convergence and wide acceptance, while providing the flexibility to accommodate the rapid technological developments in this field.

Sentence 2 – analysed in 2.1 Interpretation of the definition of an AI system using recital 12

Moreover, the definition should be based on key characteristics of AI systems that distinguish it from simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations.

Sentence 3-4 – analysed in 3.1 Learning and modelling

A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to derive models or algorithms, or both, from inputs or data.

Sentence 5-6 – analysed in 3.2 Reasoning: logic and knowledge-based approaches

The techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives, and logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved. The capacity of an AI system to infer transcends basic data processing by enabling learning, reasoning or modelling.

Sentence 7 – analysed in 2.2 Machine-based system

The term 'machine-based' refers to the fact that AI systems run on machines.

Sentence 8-9 – analyzed in 2.5 Explicit or implicit objectives

The reference to explicit or implicit objectives underscores that AI systems can operate according to explicit defined objectives or to implicit objectives. The objectives of the AI system may be different from the intended purpose of the AI system in a specific context.

Sentence 10 – analysed in 2.7 Predictions, content, recommendations or decisions and 2.8 Physical and virtual environment

For the purposes of this Regulation, environments should be understood to be the contexts in which the AI systems operate, whereas outputs generated by the AI system reflect different functions performed by AI systems and include predictions, content, recommendations or decisions.

Sentence 11 – analysed in 2.3 Varying levels of autonomy and 4. Autonomy

AI systems are designed to 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.

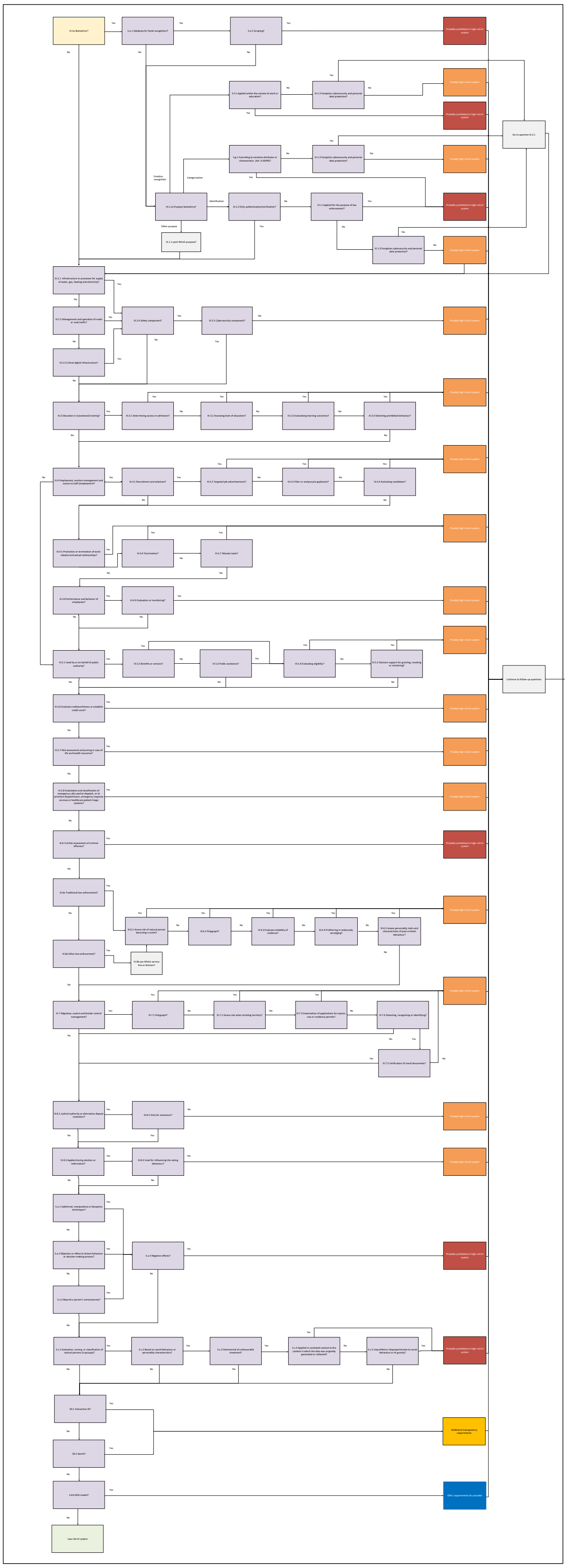
Sentence 12 – analysed in 2.4 May exhibit adaptiveness

The adaptiveness that an AI system could exhibit after deployment, refers to self-learning capabilities, allowing the system to change while in use.

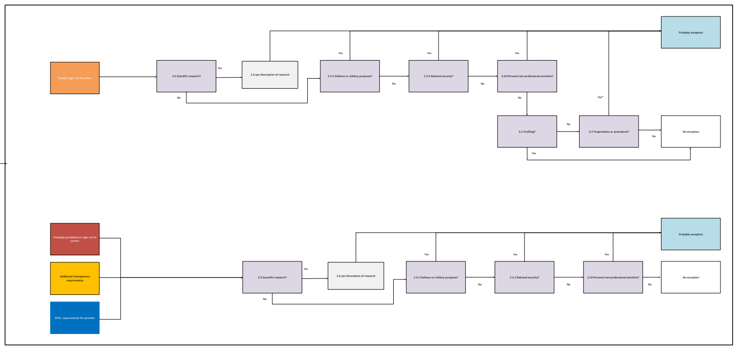
Sentence 13 – not analysed, because it does not have specific added value

AI systems can be used on a stand-alone basis or as a component of a product, irrespective of whether the system is physically integrated into the product (embedded) or serves the functionality of the product without being integrated therein (non-embedded).

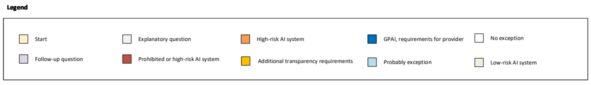
Appendix B – Flowchart



Follow-up questions exception clauses



Legend



About Algorithm Audit

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

Dutch Ministy of the Interior and Kingdom Relations

The Dutch Ministry of the Interior is committed to a solid democratic constitutional state, supported by decisive public management. The ministry promotes modern and tech-savvy digital public administrations and govermental organization that citizens can trust.

Building *AI auditing* capacity

from a *not-for-profit* perspective

www.algorithmaudit.eu

www.github.com/NGO-Algorithm-Audit

info@algorithmaudit.eu

Algorithm Audit

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