Understanding Algorithms

Pieter Oosterwijk, Miranda Pirkovski, and Berrie Zielman

1 Introduction[1]

The central government has been using algorithms for decades now. An algorithm is defined as a set of rules and instructions that a computer follows automatically when performing calculations to solve a problem or answer a question. Algorithms come in many different forms, ranging from computational models, decision trees and other statistical analyses to complex data processing models and 'self-learning' applications.

Algorithms are growing ever more popular, thanks to advancing computerisation and digitisation. Social media, navigation systems and applications like weather apps all work with algorithms. Whenever questions are asked about algorithms (for example: What is their social relevance and which risks do they pose?), the responses can be both positive and negative, in some cases extremely so. The impression arises that algorithms are becoming increasingly intelligent. This is due to the fact that, as the volume of data increases and better hardware becomes available, algorithms can process more data at greater speed, i.e. they become more innovative and wide-ranging. They can also be used for more purposes (e.g. in robotics) and, in their most sophisticated form, 'are able to correctly interpret external data, to learn from such data, and to use these learnings to achieve specific goals and tasks through flexible adaptation (Kaplan & Haenlein, 2019)'. The latter is often referred to as 'artificial intelligence' (AI). AI and algorithms are topics attracting a high level of interest from both private citizens and central government. All hold high hopes for their future potential.

[1] This chapter is based on a publication published by the Netherlands Court of Audit (2020).

P. Oosterwijk · M. Pirkovski (✉) · B. Zielman
Netherlands Court of Audit, The Hague, The Netherlands
e-mail: P.Oosterwijk@rekenkamer.nl; M.Pirkovski@rekenkamer.nl; A.Zielman@rekenkamer.nl

© The Author(s) 2023

E. Berghout et al. (eds.), *Advanced Digital Auditing*, Progress in IS,
https://doi.org/10.1007/978-3-031-11089-4_5

The wide public interest in algorithms has prompted a plethora of initiatives, standards and guidelines, developed by different stakeholders from all sorts of different perspectives. Virtually all ministries are currently working on standards and guidelines for assessing algorithms. Several non-governmental organisations are also working on the same issue, among them NOREA, the Dutch professional association of IT auditors, and large accounting firms. No comprehensive, practical tools for assessing or analysing algorithms have been developed to date, however. We take the word 'comprehensive' to mean that no efforts have been made to date to bring together all relevant standards and guidelines for algorithms into a single all-embracing framework. The word 'practical' means translating standards and guidelines into specific points that need to be assessed, the concomitant risks, and the questions that need to be answered. The audit framework forms part of this chapter and is publicly available online.[2]

In presenting this chapter, we wish to deliver a practical contribution to the debate about the opportunities and risks associated with the use of algorithms and AI in central government. The developed audit framework may provide a basis for the responsible use of algorithms and underpin the debate about the assessment and monitoring of algorithms. This chapter consists of seven sections. In the section hereafter (Sect. 2) we will provide some background and basic notions about algorithms, and how they are used in governmental practice. The third section presents the framework to audit algorithms. In Sect. 4 the case studies to test and improve the audit framework are discussed. We analysed the results of the case studies and share our main observations in Sect. 5. The main observations of the case studies are discussed in Sect. 6 while also providing several guidelines for the use of algorithms. We conclude the chapter in Sect. 7.

2 Basic Notions of Algorithms

Algorithms are shrouded in mystery and many definitions exist of what constitutes to an algorithm. We maintain the definition that an algorithm is a set of rules and instructions that a computer follows automatically when performing calculations to solve a problem or answer a question. The aim of designing an algorithm differs and depends on the task for which it is created. Several types of tasks can be discerned. There are simple algorithms that given a certain input X produces an output Y by following a well-defined set of sequential steps. This type of algorithm is predominantly used in IS to automate simple processes and is most people have in mind when thinking about an algorithm. *Descriptive* algorithms are used to describe what is happening to an output based on the input data. Sometimes the aim is also to diagnose why modifications to an output variable(s) are happening with *diagnostic* algorithms. Predictive and prescriptive algorithms are most sophisticated and have a

[2]For the full report, please visit: www.rekenkamer.nl/algoritmes-toetsingskader.

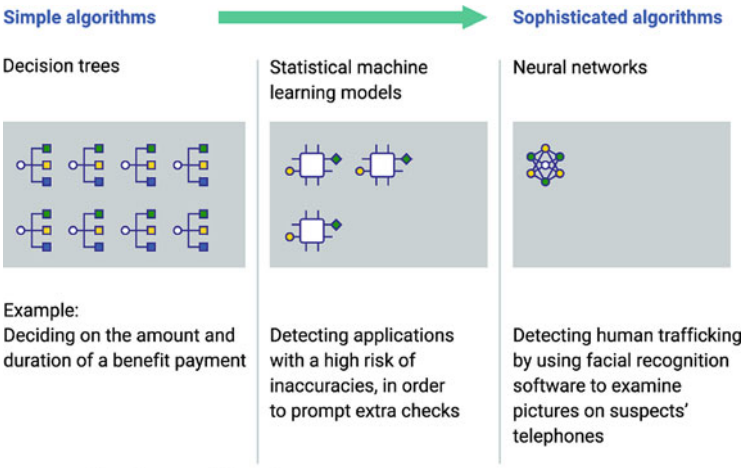

Fig. 1 Classification of algorithms

different aim. As the name suggests a *predictive* algorithm is designed to predict future outcomes based on past data. Predictive algorithms are used to answer the question 'What's going to happen next?'. Prescriptive algorithms go beyond this aim by not only calculating what is likely going to happen next, but in addition by making suggestions of what action should be taken. A prescriptive algorithm is used to answer the question 'What needs to be done?'.

Algorithms can also be classified based on complexity and explainability. In order to produce a detailed classification of algorithms, we used the information contained in the appendix to the letter of 8 October 2019 from the Minister for Legal Protection to the Dutch Parliament (Ministry of Justice and Safety, 2019a, b). The classification is based on the complexity of the algorithms, ranging from simple to complex. Figure 1 depicts the classification of the algorithms.

A *decision tree* is an example of a simple algorithm. Decisions made by such algorithms are easy to explain. An algorithm used for fixing the level of a benefit payment is a good example. A *deep-learning* algorithm, on the other hand, is a complex algorithm. Deep learning is a form of machine learning based on models similar to the neural networks of the human brain. Machine learning employs algorithms that allow computers to learn. The predictions made by this type of algorithm are difficult to analyse. It is not clear to the person making the assessment which data characteristics the algorithm regards as being more important than others. Siri (Apple's voice recognition app) and Alpha Go are two examples of such algorithms. The latter is a computer program developed by Google to play Go, a board game. In 2016, it defeated the human Go world champion.

Sitting between these two ends of the scale are algorithms of varying degrees of complexity and levels of explainability. Our analysis showed that the government uses both simple and sophisticated algorithms and both predictive and prescriptive algorithms. Most of the algorithms presented for our audit are simple algorithms and

medium-category algorithms. No more than 10% of the algorithms presented to us were categorised as sophisticated. The algorithms affect a wide range of government processes and units. A large proportion of these algorithms are used to support operating processes, thus improving efficiency. The government's use of algorithms has three purposes, each of which comes with different effects and risks. Half of the algorithms presented to us are used for the first of these purposes; the remaining half is evenly distributed over the second and third purposes.

2.1 The Use of Algorithms in Practice

We analysed the predictive and prescriptive algorithms used by the central government. This gave us an initial impression of the algorithms used in decisions affecting citizens and businesses. We asked all ministries to report the most important algorithms focusing on predictive and prescriptive algorithms. This gave us an adequate, though not comprehensive, overview of all the algorithms used by central government. We found that about one-third of the predictive and prescriptive algorithms listed by the ministries use automated decision-making. Our analysis did not identify any fully self-learning algorithms in central government, only learning ones. Automated decision-making is used only by algorithms that perform simple administrative tasks that have no effect on private citizens. Our investigation is also aimed at laying bare for what purposes algorithms are used within the Dutch government. The result of our analysis shows that the Dutch government employs algorithms for administrative activities and implementing simple legislation, to improve and facilitate operational management, and to better allocate resources based on risk predictions.

2.1.1 Automating Administrative Activities and Implementing Simple Legislation

A part of the algorithms is used to automate routine human activities. The government makes widespread use of such algorithms. This may generate big efficiency gains, particularly because they enable large volumes of data to be processed much more quickly. These algorithms often involve the (automated) implementation of legislation. A good example of one of these algorithms is the algorithm used for the listed dwellings grant scheme operated by the Cultural Heritage Agency. A decision tree (using simple 'if, then...' rules) is used to decide whether private owners of listed buildings are entitled to a grant. These algorithms are typically prescriptive and perform an automated administrative or financial activity without any human intervention. There is a low risk of errors affecting private citizens with these algorithms, as they are simple algorithms used to perform simple activities, with a high level of technical transparency and a low risk of error.

2.1.2 Improving and Facilitating Operational Management

Algorithms that are intended to boost the efficiency of government processes use more complex data. Experts cannot always blindly adopt their outcomes. These algorithms make a prediction or perform an analysis, which an expert then uses as an aid in his or her work. The Object Detection Sonar used by the Directorate-General for Public Works and Water Management is a case in point. This algorithm indicates the position of objects in the sea, based on seabed imaging, and is used to inform an expert whether it is safe to launch a hydraulic engineering project. Another example is the algorithm used to predict the number of calls made to a call centre, so that the management knows how many staff they will need. Many of these algorithms are predictive algorithms that do not involve any automated decision-making. Although there is a risk of the algorithm making errors affecting citizens or triggering a substantial level of payments, this risk is low. This is because the algorithm has only a preparatory function: it performs an analysis that an expert assesses before taking a final decision.

2.1.3 Targeted Deployment of Resources Based on Risk Predictions

The algorithms used for the third purpose are those that assist officials in selecting cases for further investigation. These algorithms help the government to deploy staff capacity and resources as efficiently as possible. The visa application process is a good example. The Ministry of Foreign Affairs uses an algorithm that helps to classify all visa applications in a number of different 'tracks'. The algorithm sorts applications into potentially successful and complex or high-risk applications, after which a governmental official checks the applications. The algorithm informs the official which applications are likely to need more time, without automatically deciding whether the application should be granted.

Previous audits have found that the central government makes widespread use of risk-based checks and our analysis confirms this. The Tax and Customs Administration does this a lot (Netherlands Court of Audit, 2019a, b, c), for example for the purpose of performing targeted audits of tax returns. The algorithm typically makes a recommendation, and it is then up to an official to decide, based on their professional judgement, whether to follow this recommendation. In other words, no automated decision-making is involved.

The algorithms supporting risk predictions carry a risk that the assumptions underlying the risk profile are not consistent with the law or may produce (undesirable) anomalies due to certain hidden limitations in the input data. The result may be a form of discrimination or the use of special category personal data. There is also a risk of the recommendation made by the algorithm influencing the official's decision.

2.2 Opportunities and Threats of Algorithms

In its Strategic Action Plan for Artificial Intelligence, submitted to the Dutch House of Representatives on 8 October 2019, the Dutch government stated that AI is a key technology (Ministry of Economic Affairs and Climate, 2019). The government is planning to invest €23.5 million in 2021 in the Dutch AI Coalition, a public-private partnership in artificial intelligence. Virtually all the ministries are either developing or already using applications. Some of these involve highly innovative algorithms using artificial intelligence. Algorithms support and in many cases improve operational management and service delivery by organisations. For instance, they enable organisations to deploy people and resources in a highly targeted way when undertaking audits or inspections. Algorithms also enable decision-making processes to be made more transparent and easier to audit. This is because the technology underlying an algorithm, the data used by the algorithm and the algorithm's interactions with these data, are all clearly defined in the form of instructions—instructions that are often absent in human decision-making processes.

In tandem with the advantages and opportunities algorithms offer, the use of algorithms by government organisations also poses several threats. The way in which an algorithm works in central government and its impact on government actions may not be sufficiently clear or may not be clearly explained to the general public. This may be related to the technology used (e.g. neural networks) or to its complexity (e.g. the algorithm may involve too many variables or components). There is also a risk that the algorithm or the data set used by the algorithm may contain certain biases that lead to discrimination. Humans also have certain in-built biases, but there is a risk in using an algorithm that it may be primarily dependent on decisions taken by the programmer or data scientist (for example, on the data used). The programmer or data scientist may lack specific knowledge and experience about the context, e.g. detailed knowledge of a decision on a grant application, even though this knowledge is essential in order to reach an informed decision. Another threat posed by algorithms that learn from data is that we often do not know or cannot foresee in advance what the algorithm will exactly learn, and to what extent there may be undesirable learning effects. Certain correlations in the data used may for instance produce an algorithm that discriminates. Finally, many algorithms used by central government have been obtained from external suppliers. This also applies to IT systems with built-in algorithms. The exact data and mechanisms used by these algorithms are often owned by the external supplier in question, who may wish to protect this information. Where liability or aspects such as the processing of personal data are concerned, the government cannot, or may not wish to, simply rely on the information provided by the supplier. This makes analysing and managing the risks associated with the algorithm more difficult for the government.

Besides being accompanied with threats and opportunities, algorithms are surrounded by myths and hypes. Algorithms are sometimes compared with human intelligence and some of them outperform humans when making certain decisions. The idea may take root that the government has lost control of its own decisions,

which may understandably lead to great unrest. When interacting with its environment, an algorithm may make a very 'intelligent' impression. However, algorithms are not intelligent. They possess neither consciousness nor sense of reality. The basic premise in the government's use of algorithms is that they should lead to greater efficiency in its operational management and the delivery of public services. Algorithms are a means to an end, and not an end in itself. Currently, most algorithms take the form of instructions that a computer follows with the help of data to reach a decision. At the same time, they are becoming both more complex and faster-acting. Combined with the potential for social unrest, this development has created a growing need among auditors and regulators for clear guidelines and assessment criteria that they can use to analyse and assess algorithms.

3 An Audit Framework for Algorithms

Algorithms bring about both opportunities and threats for governments. In this section, we present a framework to maximise the benefits algorithms have to offer while addressing potential risks. The framework was constructed by conducting an elaborate analysis of the extant literature, other frameworks, brainstorm sessions and practical analysis. A more detailed description of the methodology followed to construct the audit framework for algorithms is included in Appendix. Our audit framework contains five different perspectives for investigating algorithms that are depicted in Fig. 2. It provides concrete answers to the questions which risks are associated with algorithms, of which aspects need to be assessed.

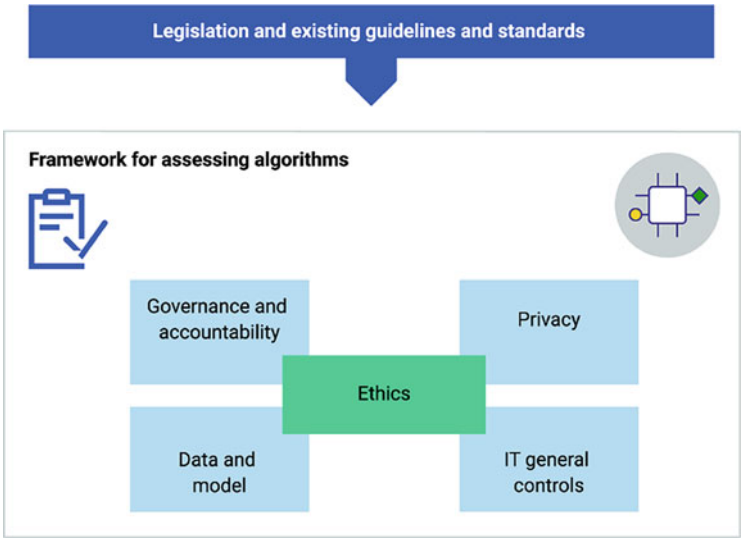

Fig. 2 Five perspectives of the framework

3.1 Ethics

Rather than forming a separate aspect of the assessment of algorithms, ethics are an integral part of the four aspects described above. In other words, ethics are relevant to all other four aspects. We identified four themes from an ethical perspective, based on existing sources (European Commission, 2020) and standards (Bergmann et al., 2019):

1. *Respect for human autonomy*—The decisions made by the algorithm are open to human checks.
2. *Prevention of damage*—The algorithm is safe and always does what it is supposed to do. Privacy is safeguarded and data protected.
3. *Fairness (a fair algorithm)*—The algorithm takes account of diversity in the population and does not discriminate. During the development of the algorithm its impact on society and the environment was taken into account.
4. *Explainability and transparency*—It is possible to explain which procedures have been followed to attain the results. It is possible to explain how the algorithm works.

3.2 Governance and Accountability

The requirements for governance and accountability focus on defining the various elements, i.e. the roles, responsibilities and expertise, the management of the algorithm's life cycle, risk factors in the use of the algorithm, and agreements with external stakeholders about aspects such as liability. We used existing IT governance standards to plan our assessment of the governance and accountability aspect of the algorithms we examined. The assessment of the governance and accountability aspect included in our audit framework is based on COBIT (Control Objectives for Information and related Technology) (ISACA, 2012) (Table 1).

3.3 Model and Data

The model and data criteria deal with questions about data quality, and the development, use and maintenance of the model underlying the algorithm. They include questions about possible biases (from an ethical perspective) in the data, data minimalization, and whether the model's output is tested. We drew on the scientific literature and the day-to-day practice of machine learning. Although the requirements we formulated as part of our audit framework focus mainly on the development of the model, they also cover operation, use and maintenance. Our audit framework is intended to cover the entire range of algorithms, from simple decision-making models to machine-learning models. Given this broadly applicable

Table 1 Risks and controls related to governance and accountability

Nr	Risk	Control
1	There can be no management or accountability without clarity about the purpose of an algorithm	The goal of the algorithm must be defined, also in relation to the social result (outcome)
2	Without an up-to-date analysis of the risks, it is impossible to reach an informed decision as to whether the benefits of using the algorithm outweigh the drawbacks	A structured and documented process for risk management
3	There is a greater risk of error without adequate resources in both qualitative and quantitative terms	An overview of the available resources (qualitative and quantitative) and management thereof
4	No full picture of the life cycle, making the algorithm impossible to manage	Lifecycle management for algorithms or the systems they are part of
5	Lack of clarity about roles, tasks, responsibilities and powers creates risks	Defined roles, described tasks, responsibilities and authorities
6	Performance and quality targets cannot be measured if there is no policy in place	An established approach to quality and performance goals for algorithms
7	A dependency on external experts who leave after developing the algorithm, taking their knowledge and experience with them, means that continuity and management are no longer safeguarded. The algorithm is not monitored and managed	Established agreements with external parties, safeguards to prevent lock-in and excessive dependence. Including exit strategy. Also consider ownership of the data used for the algorithm
8	The algorithm cannot be managed without any monitoring, leading to a higher level of risk	Organised process for monitoring the aforementioned aspects

approach this may inherently mean that certain aspects do not apply to a specific algorithm (Table 2).

3.4 Privacy

Some algorithms use personal data, including special category personal data. Sensitive data such as data revealing a person's racial or ethnic origin, religious beliefs or health status is referred to as special category data and is subject to additional legal protection (Dutch Data Protection Authority, 2022). Algorithms must comply with the statutory regulations on the processing of personal data. The General Data Protection Regulation (GDPR) is an important source of input for our audit framework (Table 3).

Table 2 Risks and controls related to a model and data

Nr	Risk	Control
1	Risk that the algorithm is not fit for purpose. Without agreement on the objectives, there is a greater risk of error and differences of interpretation	Strategic objective has been worked out in concrete terms in aspects/criteria/indicators
2	Without agreement on the objectives, there is a greater risk of error and differences of interpretation	Multidisciplinary approach and bodies
3	The operation of the algorithm cannot be explained or is difficult to explain	Explanation explicitly and, if applicable, making explicit the trade-off between explainability and performance
4	The reasons underlying the choices made in the design and implementation of the algorithm can no longer be traced (explained)	Record considerations and choices in design (such as choices between models, ROC curves) and during implementation. An ROC curve is an aid in assessing the model
5	No continuity in the process or the performance of activities, due to lack of documentation	Up-to-date, complete and accessible documentation
6	Hyper-parameters were selected at random, and the wrong choices were made in doing so	Conduct peer review (four-eyes principle)
7	A lack of transparency for private citizens, businesses and stakeholders; non-compliance with transparency legislation	Publish model (code) to a site such as github.com, including description of operation, data used and/or description thereof
8	The algorithm uses automated decision-making even though this is not permitted, or no opportunities for human intervention	Comply with applicable laws and regulations regarding automatic decision-making
9	Very limited sources of input mean a higher risk of error and non-compliance with objectives and legislation	Involve stakeholders/end users from different backgrounds in development
10	The algorithm does not operate as planned	Implementation of structural checks for correct operation
11	The model was based on the legislation applying in year $t-1$, and is now being used in year t. The legislation (e.g. on margins and limits) may have changed in the meantime, or certain legal provisions may no longer apply	Periodic checks on compliance with and in line with current laws and regulations
12	Incorrect training or testing may lead to overfitting or underfitting, or bias	Among other things, the proven separation of training, test and validation data, 'foreign eyes'/peer review and recording of process/discussions/choices
13	The model leads to undesirable systematic variance for certain individuals, groups or other units (i.e. bias)	Measures to limit, counter and/or compensate for bias
14	There is an undesirable systematic variance (bias) in the data	Check/test for bias and take countermeasures if necessary
15	A lack of separate processing leads to overfitting, which means that the model cannot be used for new observations	Visibly separated training, testing and validation data

(continued)

Table 2 (continued)

Nr	Risk	Control
16	The data is not representative	Test, check.
17	Dependency on third parties with respect to data used	Arrange for all data sources/data used that there are no restrictions/obligations
18	Violation of basic premises and rules pertaining to data minimalization and proportionality	Steering on data minimization, explicit consideration with regard to proportionality
19	The performance metrics are not consistent with the purpose of the algorithm	Good reporting/audit trail (ROC curve)
20	The data on which the model is based is available only after the outcome has been identified	Control on the mentioned aspect (target leakage)
21	The prediction meets the requisite standard	Instruments like ROC curve, confusion matrix
22	The model does not always work in practice	Monitoring output, assessing and reporting
23	People do not know that they are dealing with an algorithm. They are not aware of the consequences this has or of the algorithm's limitations. This may result in incidents, errors or claims for damages	External communication about the model/ algorithm
24	There is a risk that all efforts are concentrated on developing and producing the algorithm, and that no account is taken of the officials responsible for managing the algorithm or of the business aspects of maintenance	Maintenance and management of the technical components, the model, the data used, parameters, etc.

3.5 IT General Controls (ITGCs)

IT general controls (ITGCs) are controls adopted by organisations to ensure that their IT systems are reliable and ethically sound. These controls include conventional IT controls, such as the management of access rights, continuity and change management. The IT general controls incorporated in our audit framework focus on logging data, access rights and password management in relation to the algorithm. The requirements seek to establish whether such aspects have been built into the application and underlying components such as the database and the operating system. The main standards used for IT general controls are the international ISO/IEC 27002 standard and the Government Information Security Baseline (Table 4).

Table 3 Risks and controls related to privacy

Nr	Risk	Control
1	Not compliant with statutory regulations under the GDPR	Keeping a register according to GDPR
2	The design of the algorithm does not take sufficient account of the need to protect privacy	Design principles that ensure privacy
3	Not compliant with statutory regulations under the GDPR	Execute DPIA
4	The algorithm uses automated decision-making even though this is not permitted under the GDPR	No automatic decision-making or no documentation (for example in a privacy impact assessment) why it is allowed
5	Not compliant with statutory regulations under the GDPR; not serving mankind	Established and communicated procedure with those involved
6	Disproportionate use or collection of personal data	Recording principles, work instructions
7	Unlawful action	Recording in PIA, processing agreement/ register
8	Not compliant with GDPR or not fit for purpose	It has been established that the processing of personal data with the algorithm is compatible with the original purpose (purpose limitation)
9	Not compliant with statutory regulations under the GDPR	The lawful basis for processing personal data by the algorithm has been established
10	Violation of Article 1 of the Constitution or Article 14 of the ECHR	Think of ethnicity, skin colour, gender, sexual orientation but also zip code. Not only is checking on data itself relevant, but also so-called proxies, model bias, and so on
11	Profiling as defined in Article 4 (4) of the GDPR; risk of contravening the GDPR	Recording this review
12	Not compliant with statutory regulations under the GDPR	The data subjects are informed about the processing of personal data by the algorithm and the expected consequences
13	Not compliant with statutory regulations under the GDPR	The logic, operation and data used related to the algorithm are described and accessible
14	Not compliant with statutory regulations under the GDPR	Description and substantiation of (possibility of) human intervention in algorithm
15	Data subjects are not informed of their rights or of the algorithms and data used	There is a public privacy policy that also covers the algorithms and data used

4 Case Studies

The audit framework presented in the prior section has been submitted to a practical usability test by assessing three algorithms as case studies. Another aim of the practical usability test was to improve the framework. The aim of the practical usability test was not to arrive at any individual judgements about the algorithms, but rather to aggregate the lessons learned from the analysis. Therefore, we

Table 4 Risk mitigated through ITGCs

Nr	Risk	Control
1	Without any logging information, there is no audit trail for tracing when adjustments were made	Log information is retained and accessible until retention periods have expired. The retention period is geared to the requirements of legislation and regulations and to the control and audit cycle of the data concerned
2	Access rights are no longer up to date	Access rights are periodically reviewed and reconfirmed by the responsible management. If necessary, incidents or amendment proposals are submitted
3	Unlawful access to the algorithm	Job changes and terminations of employment are monitored for adjustment of access rights and for revocation of means of identity and authentication
4	Access rights are issued by unauthorised staff	Access rights are issued to users and administrators upon approval by an authorised officer
5	Risk of the algorithm being manipulated in cases where access rights are incompatible	Access security is implemented according to the 'nothing is allowed unless necessary' principle on all IT resources
6	The more users are granted special powers, the greater the risk of manipulation	Generic administrative accounts (root, administrator) are blocked or can only be used under registration and supervision
7	User groups are difficult to identify	Naming conventions and a system of access rights per user group and/or role apply to setting up access rights to promote maintainability of management
8	Managers and users are difficult to identify	Naming conventions are in place to identify users and administrators to aid management maintainability
9	Unclear who made changes to or worked on the algorithm	Admins perform admin work and regular user work under 2 different usernames
10	The database is open to manipulation if holders of user accounts have access to underlying components	Users have the same rights and restrictions at the application level and beyond.
11	The database is open to manipulation if holders of user accounts have access to underlying components	Job changes and terminations of employment are monitored for adjustment of access rights and for revocation of means of identity and authentication
12	The database is open to manipulation if holders of user accounts have access to underlying components	Using two-factor authentication in high-risk zones, periodically changing passwords, locking accounts when inactive, and blocking after a preset number of false login attempts.
13	Unauthorised access, changes, damage to and/or loss of data. Non-compliance with the law	Changes are tested and approved. Periodic monitoring takes place on the processed changes.

(continued)

Table 4 (continued)

Nr	Risk	Control
14	Unauthorised access, posing a risk of the algorithm being manipulated (changes, damage, loss of data)	Security
15	Back-ups are not consistent with the back-up policy. There is no recovery option, and hence a risk of data loss, if the algorithm stops working	Backup and restore policy
16	There is a much higher level of risk if there is no security by design	Security by design has been used and can be seen as the starting point. Aspects of this can be found in the ISO/IEC 27000 series and beyond.

generalised the findings of the usability test across the algorithms. A further objective was to collect more information on the risks attached to algorithms, in order to supplement the information, we had already gathered in performing our analysis. This enabled us to identify areas in which improvements are needed for the further development of algorithms in central government.

4.1 Selection of Algorithms

To test the audit framework, we selected three specific algorithms as a case study. The first algorithm is a decision tree designed to make recommendations for checks or extra checks of applications from private citizens (depicted in Fig. 3). As a second case study we selected an assessment system for detecting non-standard objects, generating information for regulators and inspectors (see Fig. 4). A facial recognition system for granting individuals physical access to a site or building was picked as the third case study. This algorithm is depicted in Fig. 5. These three predictive and/or prescriptive algorithms were selected because they are daily used, have substantial impact on both private citizens and business, and employ different techniques.

5 Analysis and Main Observations

From the analysis of the case studies, we attained some interesting observations about the use of algorithms by the Dutch government. Hereafter we will discuss these observations and their implications using the framework.

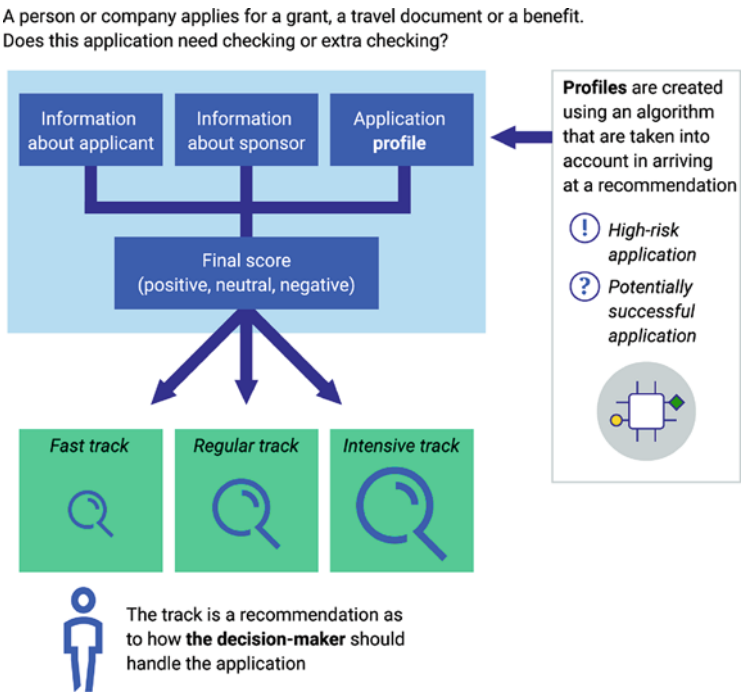

Fig. 3 Decision tree algorithm

5.1 Governance and Accountability

The extent to which the audited algorithms comply with the governance and accountability requirements differs. In the case of one algorithm, we found documentation and records extending over a number of years, explaining the basic principles and requirements applying to the algorithm. In the case of another algorithm, the documentation did not provide any clarity. This does not mean, however, that the ministry in question has no clear picture whatsoever of the purpose and operation of the algorithm. The ministry officials involved have a basic understanding of the principles underlying the algorithm. All three algorithms are subjected to regular assessments and reviews. A review means that the algorithm is reassessed in order to establish whether it still complies with the relevant standards.

In all three cases, we found that the agreements, roles and responsibilities of the parties involved in the use of algorithms in central government need to be allocated and clarified. This is necessary so that each ministry or executive agency, acting under the guidance of the CIO, can obtain a systematic understanding of whether the algorithm is doing what it is intended to do. We also found that, in many cases, no

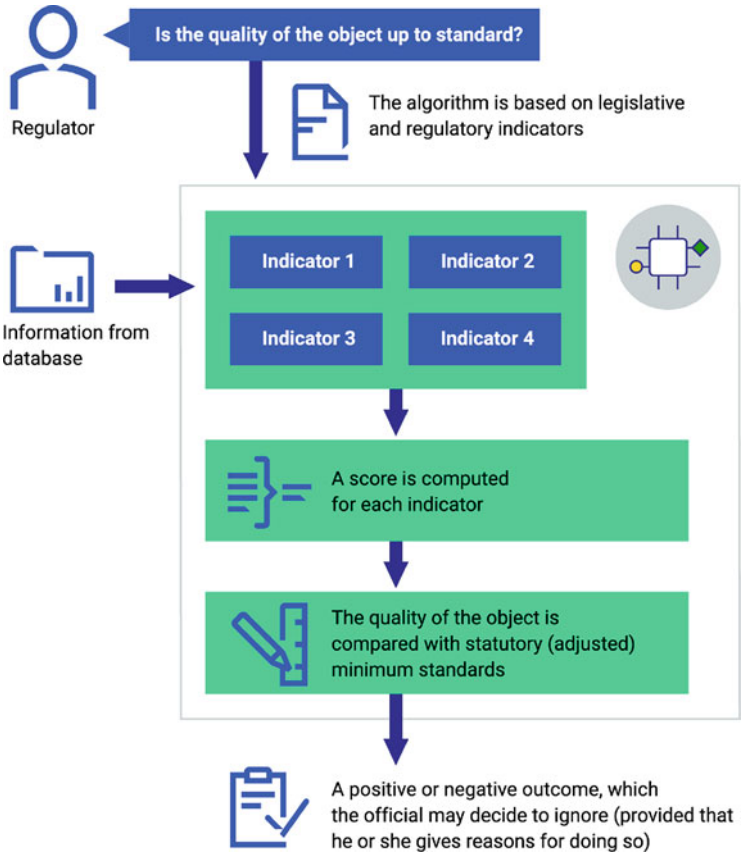

Fig. 4 Assessment system for detecting non-standard objects

system of life cycle management has been adopted for algorithms.[3] While a great deal of time and energy is spent on the design and implementation of algorithms, this does not apply to their sustainment and maintenance. This has both technical and budgetary ramifications. An inadequate maintenance budget, inadequate maintenance or inadequate staffing levels may ultimately cause the algorithm to fall short of new ethical or legal standards.

[3]The term 'life cycle management' as used in this context means the regular maintenance of algorithms during their entire life cycle, so that they remain part of a sustainable and future-proof IT landscape.

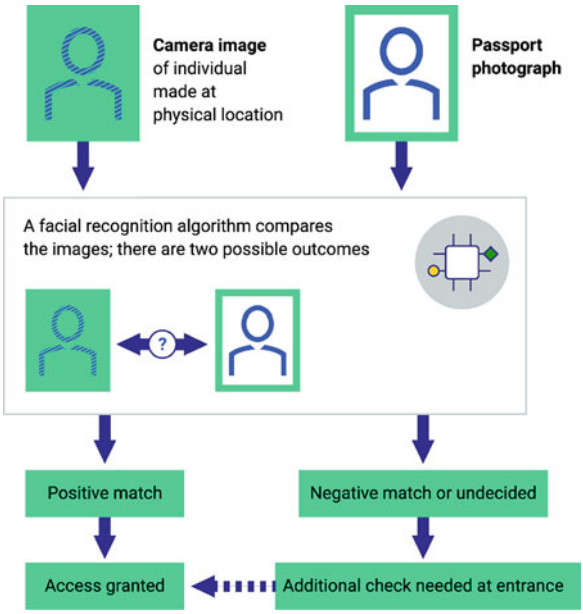

Fig. 5 A facial recognition system

5.2 Model and Data

The principle of explainability is not consistently applied. In the case of one of the three algorithms, efforts had been made to explain the model's outcome. In another case, there was a deliberate policy of avoiding transparency. The algorithm in question indicates only that there is a problem with an individual's application, without explaining why. By designing the system in this way, the executive agency wants to encourage assessors to undertake their own checks and to prevent decisions from being taken automatically without any human intervention.

The issues raised in connection with the model and data aspects include both the methods of algorithm model design and data quality. Where model design methods are concerned, we found that most officials possess sufficient expertise. There are two potential risks here in relation to data management.

The first of these is the use of historical data, which may not reflect certain social changes. This means that practices from the past are applied to the present. For instance, which competencies should a good manager possess? The answer to this question changes in accordance with social trends. If no current data is available based on new legislation, the algorithm cannot be used.

The second risk is data bias. If a specific population group was treated differently in the past, the algorithm will adopt this bias.

Our analysis of the three algorithms shows that not all relevant specialist disciplines are involved in the development of algorithms. While privacy experts, programmers or data specialists are often involved, legal experts and policy advisers tend to be left out. This may result in an algorithm failing to comply with all legal

and ethical standards or not furthering the policy objective in question. Equally, in many cases no action is taken to limit ethical risks such as biases in the selected data.

5.3 *Privacy*

The EU General Data Protection Regulation (GDPR) is the main regulatory framework for privacy and data protection. We tested the three algorithms against our audit framework. The privacy aspect involves elements such as the GDPR personal data processing register, privacy impact assessments, the legal basis for the use of data, and data minimisation. The three algorithms we assessed comply more or less fully with the privacy requirements that we believe apply to algorithms. In the case of one algorithm, the privacy policy, the data used and the algorithms were not publicly available in sufficient detail. This is important in order for third parties such as private citizens to know which data is used, how the algorithm works and how it affects them. This will become an even more important issue in the future, as the volume of data use rises, and algorithms become more complex.

As far as the algorithms we assessed are concerned, we found that there is no easy way for private citizens to obtain information about the algorithms and data used by central government. How, then, can private citizens know what impact these algorithms will have? It is not enough merely to comply with the formal requirements of the GDPR. Personal data and information submitted by private citizens belong to them, and they must know what is done with their data.

Data processing registers are not publicly available in all cases, and privacy statements linked to the algorithms we assessed are not always clear and sufficiently accessible. Although, in some cases, the operation of algorithms and the variables used have been explicitly laid down in legislation, this information is often not easy to read or understand. As a result, private citizens have only a limited understanding of algorithms. In the case of one of the algorithms we assessed, we saw that the officials involved made an extra effort to explain the variables in simple terms. This they did by translating the legislation into a list of frequently asked questions and by producing a video clip.

Building on the *Regie op Gegevens* ('Control of Data') (Dutch Government, 2022) and *MijnOverheid* ('My Government')[4] programmes, private citizens must know who they can contact with their questions about algorithms, how to notify the government about data errors, and how to object to the use of data or the outcome of algorithms. At present, Data Protection Impact Assessments (DPIAs), privacy statements and data processing registers are not sufficiently accessible and are not sufficiently clear to non-specialists.

[4] MijnOverheid is the name of a government website that members of the general public can use to receive digital messages from the government and to view their personal data.

5.4 IT General Controls (ITGCs)

It is clear from the limited amount of documentation that we received from the auditees that, of the four perspectives of our audit framework, it is the ITGC requirements that are given the lowest priority. The main functions addressed by ITGC are access rights and their management, and back-ups. In two of the three algorithms we assessed, little or no information was available as to whether the relevant ITGC standards were met,[5] and auditees were either unable to provide this information or unable to provide it at short notice. In the case of the third algorithm, we did receive the documentation we requested after providing a further explanation. In conclusion, two of the three algorithm owners were unable to provide sufficient proof that they are in sufficient control of the relevant risks. We believe there are two reasons for this.

The algorithm is managed by an external service-provider. Although the relevant officials assume that these external service-providers have proper IT controls, they do not know whether this is actually the case. When we asked for proof, the officials at the ministry in question were unable to provide it or were unable to provide it at short notice.

Although the organisation in question has set higher or different ITGC standards, these have not been laid down in sufficient detail for the algorithm in question.

Our government-wide analysis of algorithms confirms the existence of the first cause, i.e. that the management of algorithms has been outsourced to external suppliers. This applies to two of the three algorithms in our practical test. In the case of one of these, a public-sector shared service organisation (SSO) had been made responsible for managing the algorithm. In the second case, the algorithm was managed by an external service-provider.

As a result, we were unable to establish whether the algorithms complied with a large number of ITGC standards. In the case of the algorithm managed in-house by a ministry, the officials concerned were able to provide documentation on all perspectives of our audit.

5.5 Ethics

Rather than forming a separate aspect of the assessment of algorithms, ethics are an integral part of the four aspects described above. We analysed each use case based on the ethical principles that underpin the framework (see Sect. 3.1).

[5] The relevant standard here is the Dutch Government Information Security Baseline, based on the international ISO/IEC 27002 standard.

5.5.1 Respect for Human Autonomy

Our audit showed that the three algorithms work as an assistive resource; they do not (or do not yet) take any automated decisions. In one case, the technical application (i.e. the algorithm) allows officials to consult several different sources, thus enabling them to take efficient decisions. In other words, the algorithm assists officials.

5.5.2 The Prevention of Damage

In order to prevent any damage, it is vitally important that the algorithm should always do what it is supposed to do. In addition, people's privacy must be safeguarded, and the relevant data must be protected. Unauthorised access may lead to data being changed, damaged or lost. Our findings are explained under the heading ITGG.

5.5.3 Fairness

Fairness means that the algorithm takes account of population diversity and does not discriminate. If no effective measures are taken, the algorithm may acquire an undesirable systematic bias in relation to certain individuals, groups or other entities. In the case of one of the three algorithms we assessed, an external supplier tested the algorithm for any undesirable outcomes. In another case, an external supplier tests all data in advance, in order to assess whether it is absolutely necessary for the algorithm to fulfil its purpose.

5.5.4 Explainability and Transparency

Owners of algorithms are obliged to explain how they designed the algorithm and how it works. All three algorithms were explainable and in all three cases the model designers sought to strike a balance between explainability and performance. Self-learning algorithms were not involved in any of the three cases, and this is one of the factors that make the algorithms in question relatively easy to explain.

In order for procedures to be explained, they need to be clearly documented. We found that this was an issue both in the case of algorithms managed in-house and in the case of those that are fully managed by external suppliers. In the former case, the parameters had been documented, but the model design had not.

In order to assess whether an algorithm adheres to the ethical principles of fairness, explainability and transparency, independent assessors must be able to identify and check the data used. In the case of one algorithm, the data needed to comply with privacy legislation was not stored. This means that, as independent assessors, we were unable to check the data after the algorithm was run (although an

external service-provider did check the data before the algorithm was run). As a result, while the algorithm does comply with privacy legislation, we were unable to establish whether the ethical principles were observed.

6 Discussion

The main observations we derived from our analysis raise some interesting points for discussion. In this section, these points will be discussed and some guidelines are proposed to control the use of algorithms.

6.1 An Algorithm Does Not Have to Be a Black Box

Algorithms are used to support human actions. Our analysis of algorithms used in central government did not reveal the existence of any algorithms that act fully autonomously. We did find algorithms that take simple decisions or perform routine activities in a non-complex environment. Automatically generated letters and messages are examples of such algorithms. Choices about explainability and transparency are part and parcel of the process of developing algorithms. Accountability is another aspect choices. If priorities are given to these aspects in the development of an algorithm, it does not become a black box, but instead a means of assisting an operating process. It should be clear which data it uses, how the data model works, which outcomes it delivers and what sort of impact these outcomes have. It should be possible to make it easier to verify the outcomes of an algorithm than would be the case with the results of a human analysis. Algorithms obtained from private suppliers are a potential problem here. They must comply with the same requirements as those developed by the government itself.

6.2 No Insight Information: Need for Specific Tools

Algorithms are often developed from the bottom up, i.e. on the basis of day-to-day working practices. Senior ministry officials and Chief Information Officers (CIOs) at ministries have little insight in this process. As a result, ministers are unable to mitigate the potential adverse effects of algorithms on government service delivery in a timely manner. The analysis in this audit should help ministers to gain a clearer picture of the way in which algorithms are used by their ministries. A further problem is that there is no standardised terminology in relation to algorithms. This accounts for our finding that ministry officials use different definitions of algorithms and different terms in describing how algorithms are developed, the associated risks and the means of mitigating these risks.

The assessment frameworks in current use are inadequate for the purpose of assessing algorithms. Ministries use universal standards such as the General Data Protection Regulation (GDPR), the Government Information Security Baseline, the Information Technology Infrastructure Library (ITIL) (ITIL Foundation, 2019) and COBIT (ISACA, 2012) for improving the quality and reliability of algorithms and for mitigating the risks attached to their use. This does not apply to all ministries, however. Ministries also use letters to the House of Representatives about big data and algorithms as guidance.

Officials from just three ministries told us explicitly that they regarded ethical aspects as an important component of algorithms. This finding is confirmed by the outcome of our practical test, in which we generally found that no action had been taken to curtail biases (e.g. in the data selection and the risk of discrimination) and a lack of attention for ethical aspects such as profiling. The general standards frameworks do not apply specifically to algorithms and are not used as an interconnected whole. Without any adequate management of and accountability for algorithms, it is impossible to make a clear analysis of the pros and cons of their use. Moreover, the effects of an algorithm are difficult to explain. They may have a significant impact on private citizens in the form of discrimination, inaccurate profiling or financial implications.

Ministry officials all agree that there is a need for a set of standards containing clear, practical definitions of algorithms. At present, there are often differences of interpretation. Opinions differ on whether these definitions should be specific or generic. Some officials regard algorithms as IT tools to which the same generic standards could apply. Other officials claim that the risks attached to algorithms are not always generic, which means that a single, generic set of standards would be impractical. The results of our brainstorming session confirm these findings.

Observation 1: publish clear, consistent definitions and quality requirements.

We urge the cabinet to adopt a clear, uniform set of terms and specific quality requirements for algorithms. Clear, consistent definitions and quality requirements will foster knowledge sharing, streamline processes and prevent misinterpretations. The officials participating in our brainstorming session provided more detailed information about this need for clear, consistent definitions in central government, and in doing so laid the foundations for a ‘common language’ for algorithms. We organised this brainstorming session in conjunction with the Ministry of the Interior and Kingdom Relations, the Ministry of Justice and Security, and the Radiocommunications Agency of the Ministry of Economic Affairs and Climate Policy. The brainstorming session presented these organisations—as pioneers in the use of algorithms in central government—with an opportunity to formulate clear, broadly applicable guidelines and quality requirements for algorithms.

6.3 Predictive and Prescriptive Algorithms Still Under Development: Limited Impact on Private Citizens to Date

Our analysis has shown that central government makes widespread use of both simple and complex algorithms. Broadly speaking, algorithms are used for three purposes:

1. For automating administrative work and simple legislation.
2. For facilitating and improving operational management and/or service delivery.
3. For performing risk-based checks and ensuring that staff and resources are deployed in a targeted manner.

We did not find any fully self-learning algorithms in central government, only learning ones. Only those algorithms that perform simple administrative activities with no substantial impact on private citizens take automated decisions.

6.4 Insufficient Account Taken of Private Citizens

Currently, Data Protection Impact Assessments (DPIAs), privacy statements and data processing registers are not sufficiently accessible and are not sufficiently clear to non-specialists and non-professionals. Private citizens do not know who they can contact their questions about algorithms, how to notify the government about data errors, and how to object to the use of data or the outcome of algorithms. In our opinion, it does not suffice merely to comply with the formal requirements of the GDPR, as this does not generally provide citizens with sufficient information about the algorithms that affect them. Central government can prevent prejudices about algorithms from arising by communicating transparently about the use of algorithms, about the effects they may have on private citizens, and about its own accountability.

Observation 2: inform private citizens about algorithms and explain how they can obtain further information about them.

We urge the cabinet to enable private citizens to access, in a logical location, information on which data is used in which algorithms, how these algorithms basically work and what impact their outcomes have. The algorithms involved here would be those that have a substantial impact on government behaviour or on decisions relating to specific cases, individuals or businesses. One option would be to create a dashboard similar to that created to provide information about large IT projects.

6.5 Improvements for the Responsible Use and Refinement of Algorithms

6.5.1 Governance and Accountability

We found that the agreements, roles, tasks and responsibilities of the parties involved in the use of algorithms in central government need to be further defined and clarified. This is necessary in order to allow ministries to obtain a systematic understanding of whether an algorithm is doing what it is supposed to do. This applies especially to cases in which multiple parties are involved in the development, operation and maintenance of the algorithm. We want to draw attention to the quality of testing of algorithms and continuous monitoring by the ministry.

We found that, in many cases, no system of life cycle management has been adopted for algorithms. While a great deal of time and energy is spent on the design and implementation of algorithms, this does not apply to their sustainment and maintenance. This may ultimately cause the algorithm to fall short of new ethical or legal standards, for instance, or simply to become technically obsolete.

Observation 3: document agreements on the use of algorithms and make effective arrangements for monitoring compliance on an ongoing basis.

Our recommendation to the cabinet is to ensure adequate documentation of the terms of reference, organisation, monitoring (e.g. in terms of life cycle management: maintenance and compliance with current legislation) and evaluation of the algorithm, as this makes clear whether the algorithm is and remains fit for purpose. This also enables the algorithm to be adjusted, if necessary. Especially if algorithms are outsourced or purchased from another (outside) supplier, it is important to ensure that all arrangements relating to liability are laid down in a contract. Our audit framework contains a number of key requirements that can be used as input for documenting such agreements.

6.5.2 Model and Data

Central government uses algorithms ranging from simple decision trees to complex algorithms for image analysis in a wide range of areas. This means that not all the aspects of our audit framework apply to each algorithm. Context also plays an important role in assessing the findings about an algorithm. While explainability may be an important means of providing citizens with information in one particular case, the same level of explainability may be undesirable in another situation, as this would influence decision-makers too much. Moreover, transparency might actually encourage fraudulent behaviour on the part of private citizens. Our audit framework can be refined into a set of standards or minimum quality requirements for any given algorithm.

The issues raised in connection with the model and data aspects include both the methods of algorithm model design and data quality. Where model design methods

are concerned, we found that most officials possess sufficient expertise. There are two potential risks here in relation to data management. The first of these is that the use of historical data may not reflect certain social changes. This means that practices from the past are applied to the present. The second risk is data bias. If a specific population group was treated differently in the past, the algorithm will adopt this bias.

Our analysis of the three algorithms shows that not all relevant specialist disciplines are involved in the development of algorithms. If legal experts and ethical specialists are not consulted, this may result in an algorithm failing to comply with all legal and ethical standards or not furthering the policy objective in question. Equally, in many cases no action is taken to limit bias (for example, in data selection or a risk of discrimination) and ethical risks.

Observation 4: ensure that the audit framework is translated into practical quality requirements for algorithms.

We recommend that the cabinet instructs the Minister of the Interior and Kingdom Relations to ensure that the Chief Information Officer at each ministry is made responsible for translating the audit framework (which is designed to assess algorithms already in use) into a practical set of design standards or into quality requirements for the development of algorithms. The objective here would be to ensure that quality requirements are more practical and could already be applied during the development stage of an algorithm.

Observation 5: ensure that all relevant disciplines are involved in the development of algorithms.

Our recommendation to the cabinet is to involve all relevant disciplines and types of specialist expertise in the development of algorithms. This means involving legal experts, ethical specialists and policy advisers alongside technical specialists.

6.5.3 Privacy

There is no easy way for citizens to obtain information on the privacy guarantees applying to the use of algorithms. This translates into the following practical issues:

Merely complying with the formal requirements of the GDPR is not an adequate means of informing private citizens about how algorithms work, the data they use and their impact.

The government's online data processing register[6] gives readers the impression that it contains all processing registers. This is not the case, however. Nor is there any legal obligation for all processing registers to be published on this website. Our recommendation for privacy is included in Sect. 6.4.

[6]For the register, please visit: www.avgregisterrijksoverheid.nl.

6.5.4 IT General Controls (ITGCs)

In those cases, in which the management of an algorithm has been outsourced to an external supplier, we found that official working with algorithms do not know whether adequate ITGCs have been put in place. Although this is not a problem in itself, we do see certain risks in the current arrangements made for the algorithms we assessed.

Ministries that have outsourced the development and management of algorithms have only a limited knowledge of these algorithms. The outsourcing ministry assumes that the supplier is in control and complies with the ITGC and other standards included in our assessment. We found no proof of this: the responsible minister does not have any information on the quality of the algorithm in question nor on the documents underlying compliance with the relevant standards and refers to the supplier instead.

Where ministries have outsourced the management of algorithms to a public-sector shared service organisation, the situation is the same as where management is outsourced to an external contractor. The department using the algorithm refers to the ITGC guidelines applying at a higher or different level of the organisation. In other words, while disclaiming responsibility, the officials at the ministry using the algorithm cannot explain how the organisation-wide standards apply to the specific algorithm in question.

Observation 6: ensure that clear information is produced now and in the future on the operation of IT General Controls.

We recommend that the cabinet instructs the Minister of the Interior and Kingdom Relations to ensure that the relevant ministers and state secretaries see to it that officials working with algorithms have and retain access to information on the quality of the ITGCs in relation to the algorithms in question. This they can do by asking the party managing the algorithm to present formal statements, such as IT auditors' reports, showing that the ITGCs are of an adequate standard.

6.5.5 Ethics

We found that legislation is sometimes inconsistent with ethical standards. In order to assess whether an algorithm adheres to the ethical principles of fairness, explainability and transparency, independent assessors must be able to identify and check the data used. The demands of privacy legislation mean that a large volume of data is not kept for very long, making it impossible for an auditor to audit it in retrospect. Independent auditors would already like to see an amendment made to the privacy law applying to complex algorithms, and this need is only likely to increase as algorithms grow more complex. This will become clear from the way in which algorithms develop in the coming years.

7 Conclusions

The audit framework that is presented in this chapter makes maximum use of existing information, guidelines and standards. Our audit framework is a practical tool that we intend to use in our future audits. Other government organisations are also free to use our framework to assess whether their own algorithms meet certain quality standards, and whether the risks are sufficiently clear and/or are being mitigated. We hope to have been clear and transparent about any questions that may arise in future audits of algorithms. Our audit framework already gives the ministries a good idea of the risks that we have identified, which means that they can start taking action to mitigate these risks now. The audit framework enables auditors to analyse algorithms from five perspectives:

- Ethics.
- Governance and Accountability.
- Model and Data.
- Privacy.
- IT General Controls (ITGCs).

We investigated how algorithms work in practice in central government and identified potential improvements. Questions about algorithms—what they can do and what risks do they pose?—elicit a wide range of reactions, ranging from extremely negative to extremely positive and everything in between. The audit framework we developed may serve both as a basis for the responsible use of algorithms and as a starting point for discussions on how to manage and monitor algorithms. Our intention is to promote transparency and to foster an open debate about the potential risks arising from the use of algorithms. Transparency about algorithms and control of their operation must become the rule rather than the exception.

Our main conclusion based on the algorithms we analysed in Sect. 4 is that central government pays a great deal of attention to mitigating the privacy risks at play in the use of algorithms. We found automated decision-making only in algorithms performing simple administrative activities that have no impact on private citizens. We also found that the complex algorithms that we analysed do not take independent decisions. Government officials play a prominent role in the use of these algorithms, which assist them in performing analyses and taking decisions.

We also found that algorithms are not a black box for us as independent auditors: we were able to examine and assess them. This does not detract from the fact that there is still room for improvement in 2021, as the use of algorithms is set to increase in the coming years. If algorithms become self-learning, i.e. more complex, they will produce better decisions in terms of speed, quality and objectivity. This will put officials at a greater distance from government decisions on private citizens and businesses. This chapter presents our conclusions and recommendations.

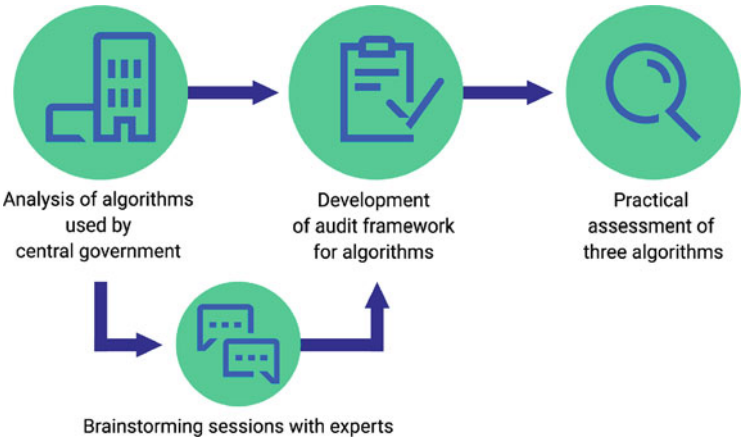

Fig. 6 Method used to construct the framework

Appendix: Methodology of the Audit

We performed an exploratory assessment of predictive and prescriptive algorithms that have a relevant impact on the operating processes of and/or service provision by central government and its associated organisations. This audit was premised on the following audit questions:

1. For which activities and processes do central government and its associated organisations use algorithms, which types or categories of algorithms are there, and what are the risks and effects associated with the use of algorithms?
2. How do the central government and its associated organisations manage the operation and control the quality of algorithms?

In order to answer these questions, and to construct the framework we followed the method depicted in Fig. 6.

Analysis of Existing Algorithms

As a first step, we analysed the types of algorithms used by central government and the activities for which they are used. Our audit builds on the classification described in the appendix to the letter to Parliament about the safeguards against the risks posed by data analysis performed by government (Ministry of Justice and Safety, 2019a, b). The appendix also differentiates between the way in which algorithms are used and the impact that they have. The impact ranges from small in the case of descriptive algorithms to big in the case of prescriptive algorithms.

We asked the ministries to submit examples of prescriptive and predictive algorithms with a relevant impact on the government's operating processes and/or

service delivery. We asked ministries for their most representative algorithms. There was space in the questionnaire for ten algorithms, but this was merely an indicative number. For the purpose of this audit, we wished to receive information about algorithms that have both: (1) a *predictive* or *prescriptive* function, and (2) a substantial impact on government behaviour, or on decisions made about specific cases, citizens or businesses. We looked at the purposes for which these algorithms are used, the impact that they have on citizens, and how they are managed and documented.

As the focus of our audit lies on substantial impact, we elected to analyse predictive and prescriptive algorithms. We wish to stress that we did not seek to undertake a comprehensive analysis of all the algorithms used by central government. We asked the ministries to self-report on the algorithms they used which they believed met our specifications. We explored certain issues in more detail during interviews. We drew up reports of the interviews, which we then asked the interviewees to check.

Brainstorming Session in September 2020

During the course of our analysis, it became clear to us that operational staff responsible for the design, implementation and management of algorithms wished to see closer cooperation among the ministries and needed practical tools for using algorithms in a responsible manner. In order to meet these needs, we organised a brainstorming session on 22 September 2020 in conjunction with the Ministry of the Interior and Kingdom Relations, the Ministry of Justice and Security, and the Radiocommunications Agency of the Ministry of Economic Affairs and Climate Policy. These organisations are pioneering the use of algorithms in central government. Thirty experts from both within and beyond central government took part in the session.[7]

When it became clear during the course of our research that all the stakeholders involved in the use of algorithms worked with different definitions of algorithm-related terminology, we organised a brainstorming session on 22 September 2020. We did this in conjunction with the Ministry of the Interior and Kingdom Relations, the Ministry of Justice and Security, and the Radiocommunications Agency of the Ministry of Economic Affairs and Climate Policy. The aim of the brainstorming session was to identify, discuss, and, if possible, bridge the differences in the terminology used for algorithms. The brainstorming session was broken down into five themes:

- Data-driven
- Data quality

[7] In compliance with Covid-19 restrictions, only a small number of experts were allowed to attend the brainstorming session.

- Artificial intelligence and algorithms
- Artificial intelligence in central government
- Transparency

Constructing the Audit Framework

The audit framework that we used for this audit is based on various types of existing information, parameters and standards. Our audit framework is a practical tool that we intend to use in future audits. However, other government and private-sector organisations are also free to use it to assess whether their algorithms meet specified quality criteria, and whether the accessory risks have been properly identified. The audit framework is a component part of this report and is publicly accessible at: www.rekenkamer.nl/algoritmes-toetsingskader.

Practical Assessment of Three Algorithms

Subsequently, we selected three algorithms from our list and tested them with the help of our audit framework. Our purpose was to refine our audit framework by submitting it to a practical test. By assessing algorithms we can identify those areas where improvements are required in how the central government manages the risks relating to its use of algorithms.

We analysed the predictive and prescriptive algorithms used by the central government. This gave us an initial impression of the algorithms used in decisions affecting citizens and businesses. We asked all ministries to report the most important algorithms focusing on predictive and prescriptive algorithms. This gave us an adequate, though not comprehensive, overview of all the algorithms used by central government.

We found that about one-third of the predictive and prescriptive algorithms listed by the ministries use automated decision-making. Our analysis did not identify any fully self-learning algorithms in central government, only learning ones. Automated decision-making is used only by algorithms that perform simple administrative tasks that have no effect on private citizens.

The ministries' responses show that, with the exception of the Ministry of General Affairs (which does not use any algorithms that are within the scope of this research), they all use both predictive and prescriptive algorithms for delivering services (depicted in Fig. 7). The ratio of predictive to prescriptive algorithms is virtually the same: 60% of the algorithms used are predictive.

The number of predictive and prescriptive algorithms submitted for the purpose of this audit differs from one organisation to another. Large organisations such as the Employee Insurance Agency and the Social Insurance Bank distribute funds, benefits and grants in accordance with statutory regulations. These institutions typically use prescriptive algorithms. The number of algorithms used is not necessarily a

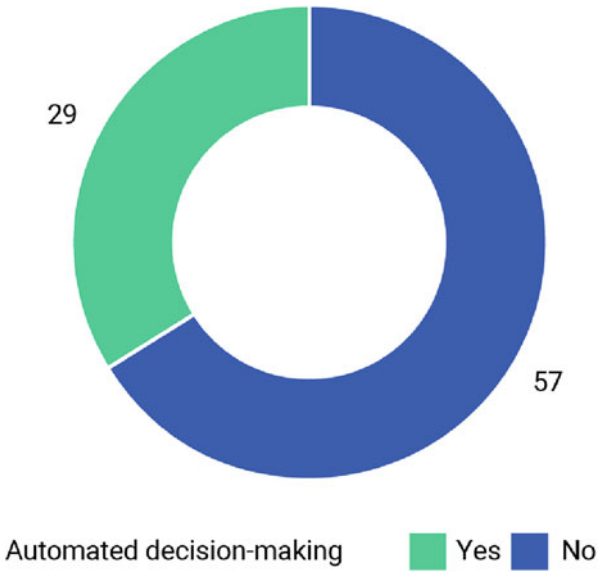

Fig. 7 Overview of the types of algorithms used by the Dutch government

reflection of the degree of expertise on algorithms that a given organisation possesses, as they differ in terms of their complexity and potential impact. We also found that central government does not have any uniform definition or standardised classification of algorithms, which resulted in differences of interpretation among the ministries when submitting their algorithms.

Virtually all the ministries, as well as the central government CIO, informed us that they have no comprehensive, centralised list or overview (i.e. maintained by the ministry itself) of the algorithms used by the ministry in question. As a result, the ministers are unable to timely mitigate the risks and potential adverse effects of algorithms on government services. The same lack of overview also applies to organisations associated with ministries (see the figure above). A number of ministries and the central government CIO told us that our audit was the first step towards obtaining a realistic picture of their use of algorithms.

References

Bergmann, U., Bonefeld-Dahl, C., Dignum, V., Gagné, J.-F., Metzinger, T., Petit, N., et al. (2019). *Ethics guidelines for trustworthy AI*. European Commission. Retrieved from https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf

Dutch Data Protection Authority. (2022, June 3). *Wat zijn persoonsgegevens?* Retrieved from Dutch Data Protection Authority: https://autoriteitpersoonsgegevens.nl/nl/over-privacy/persoonsgegevens/wat-zijn-persoonsgegevens

Dutch Government. (2022, June 3). *Digital Government*. Retrieved from Control of Data: https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/regie-op-gegevens/

European Commission. (2020). *Whitepaper on artificial intelligence—A European approach to excellence and trust. European Commission*. European Union.
ISACA. (2012). *COBIT 5, a business framework for the governance and management of enterprise IT*. ISACA.
ITIL Foundation. (2019). Axelos.
Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons, 62*(1), 15–25. Retrieved from https://www.sciencedirect.com/science/article/pii/S0007681318301393
Ministry of Economic Affairs and Climate. (2019). *Strategic action plan for artificial intelligence*. Ministry of Economic Affairs and Climate.
Ministry of Justice and Safety. (2019a). *Waarborgen tegen risico's van data-analyses door de overheid*. The Hague.
Ministry of Justice and Safety. (2019b). *Waarborgen tegen risico's van data-analyses door de overheid*. Dutch Government. Retrieved from https://www.tweedekamer.nl/kamerstukken/detail?id=2019Z19084&did=2019D39751
Netherlands Court of Audit. (2019a). *Cyber security and critical water structures*. Netherlands Court of Audit. Retrieved from file:///C:/Users/bjbut/Downloads/Vertaling+Cybersecurity+WR.pdf.
Netherlands Court of Audit. (2019b). *Data-driven selection of tax returns by the Dutch Tax and Customs Administration*. Netherlands Court of Audit. Retrieved June 11, 2019, from https://english.rekenkamer.nl/publications/reports/2019/06/11/%E2%80%A2data-driven-selection-of-tax-returns-by-the-dutch-tax-and-customs-administration
Netherlands Court of Audit. (2019c). *Informatiebeveiliging Verantwoordingsonderzoek*. Netherlands Court of Audit.
Netherlands Court of Audit. (2020). *Cyber security of border controls operated by Dutch border guards at Amsterdam Schiphol Airport*. Netherlands Court of Audit. Retrieved from file:///C:/Users/bjbut/Downloads/Cybersecurity+Engels+WR.pdf.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.