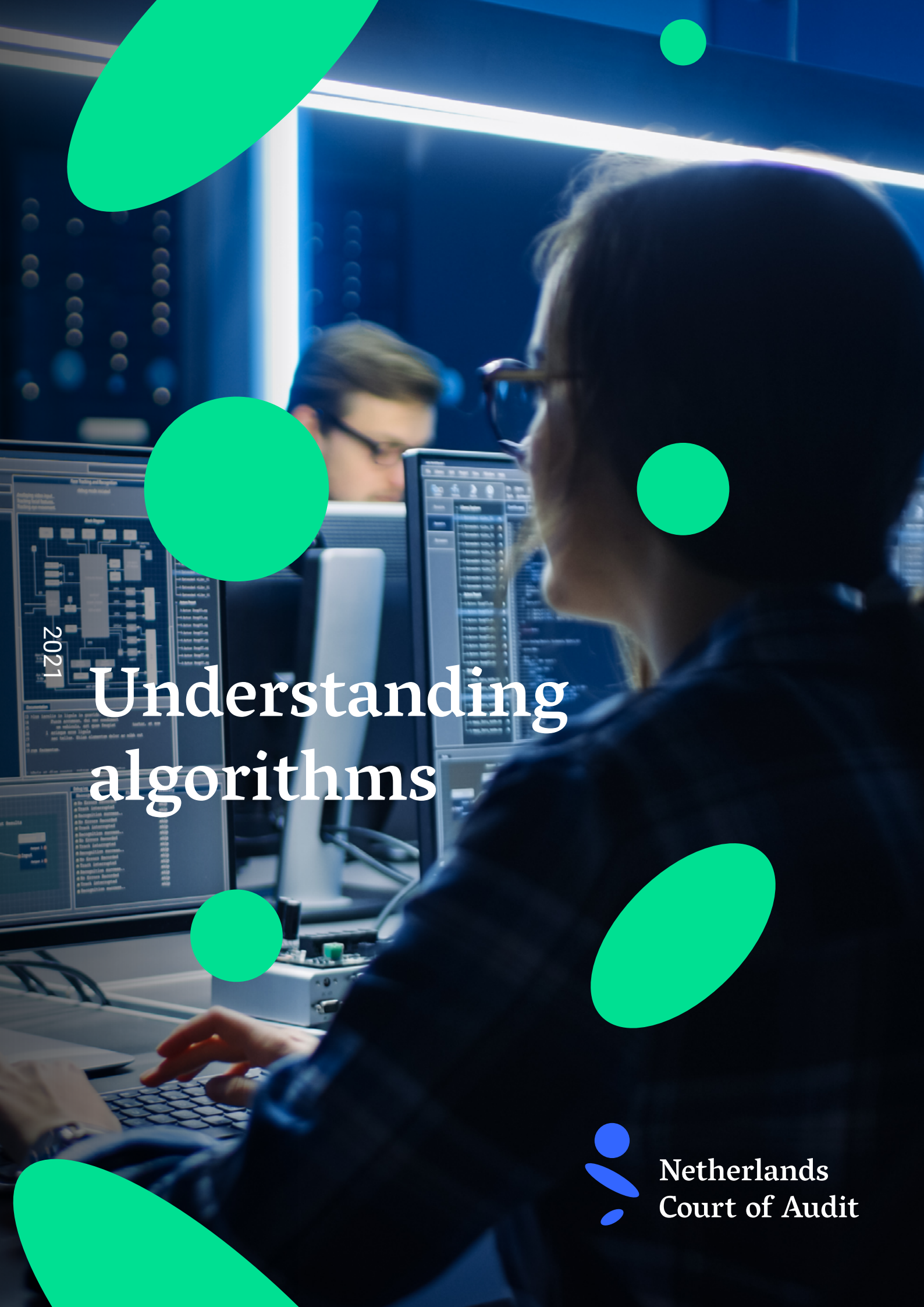


2021

Understanding algorithms

Netherlands Court of Audit

Preface

We launched this audit in early 2020, at a time when the Dutch government had just announced measures to contain the outbreak of Covid-19 and was forced to devote all of its attention to managing the crisis. Our audit period coincided exactly with the first wave of Covid-19 infections. The second wave of infections struck the Netherlands at the point when we were discussing our audit findings with ministry officials. Despite the impact that Covid-19 has had (and is still having) on our daily lives, ministry officials and their departments supplied us with all the information that we asked for and also made time available for answering our questions during interviews. Partly thanks to their efforts, we were able to continue our audit in these exceptional circumstances.

Original title
Algemene Rekenkamer (2021). *Aandacht voor algoritmes.*

Contents

1. Summary

Central government uses algorithms in implementing its policies. Algorithms are sets of rules and instructions that a computer follows automatically when making calculations to solve a problem or answer a question.[1] We wanted to demystify these algorithms by finding out what they actually do and what they don't do. We intended to answer questions such as: How does the government avoid bias when it is using algorithms? Does the government oversee the consequences that the use of algorithms has on private individuals and companies that are affected by government policies?

Take for instance System Risk Indication (SyRI). This is an algorithm-based system used by central government (e.g. by the Employee Insurance Agency and the Tax and Customs Administration) to detect fraud. In February 2020, the Court of Justice ruled that the legislation regulating the use of SyRI represented an unacceptable infringement of citizens' privacy rights.[2]

Dutch Members of Parliament have also regularly expressed their concerns about discrimination and biases, which they claim are a constant risk associated with the use of algorithms. The public debate in mid-2020 about the Covid-19 notification app (CoronaMelder) came on top of these concerns. In addition to focusing on issues relating to the source, collection and use of data, the debate also centred on the transparent and verifiable operation of the used algorithms.

Algorithms account for an ever larger component of the central government's operations and actions, and hence perform an ever important role in the delivery of public services to citizens and businesses. We analysed the activities and processes for which central government and its associated organisations use algorithms, classified these into categories, and identified the risks involved in the use of algorithms. In addition, we examined how central government and its associated organisations manage the operation and control the quality of algorithms.

1.1 Conclusions

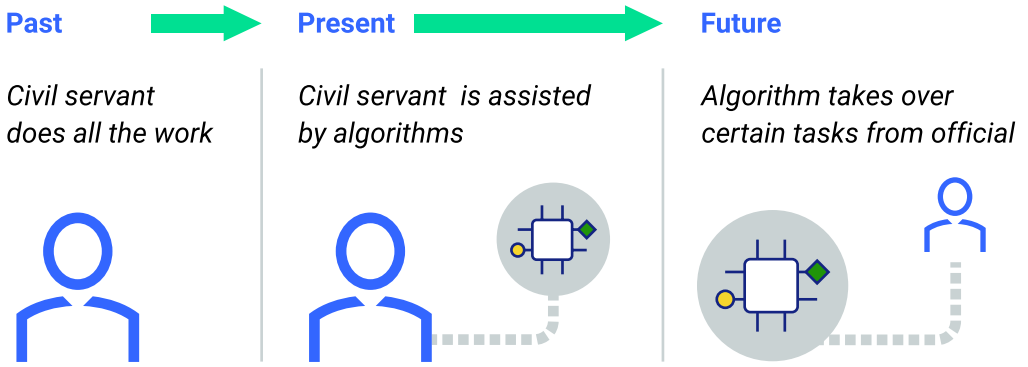
We found that most of the algorithms used by central government are relatively simple. They have a limited effect on private individuals, as it is only such relatively simple algorithms that take automatic decisions. Many of these decisions involve the automation of certain administrative activities, for example the automated sending of letters confirming the receipt of a communication. We did not find any fully self-learning algorithms in central government, only learning algorithms. Humans are always involved in the algorithm's learning process. In other words, there is always a 'human in the loop'.

Our audit shows that algorithms are not a black box for us as an independent auditor: we were able to examine and assess all of the algorithms we identified. We also found that, based on the predictive and prescriptive algorithms[3] we analysed, the government devotes a great deal of attention to containing privacy risks in the development and use of algorithms. We also found that the algorithms we analysed do not take decisions independently. Operational staff are explicitly involved in the use of these algorithms. The algorithms assist the staff concerned in making analyses and taking decisions.

This does not detract from the fact that – viewed from the aspect in the year 2021 – there is room for improvement, as algorithms are set to be used more and more often in the years ahead. If algorithms become self-learning,[4] and hence more complex, they will produce better decisions in terms of speed, quality and objectivity.

This will also have the effect of operational staff being less involved with regard to decisions impacting citizens and businesses. If this is the case, the quality of algorithms will need to meet stricter standards. This is why it is important that the cabinet – in the first instance, the Minister of the Interior and Kingdom Relations – wastes no time in addressing the issues and recommendations raised in this report. We would also like to stress that complying with standards in relation to cyber security and data protection is a crucial precondition for the responsible use of algorithms. The challenges here include preventing and detecting cyber attacks such as digital sabotage, espionage and cyber crime.[5]

Government officials are making more use of algorithms



Despite the widespread public interest in algorithms, no specific tools for auditing or analysing algorithms have not yet been developed / up untill this date. This is why we developed our own audit framework. It incorporates the standards that are currently used in order to limit the potential risks inherent to the use of algorithms. We link the aspects that are tested and the audit questions to these risks. The likelihood of the risks actually materialising in relation to a specific algorithm, and the extent of the damage thus caused, depend on whether or not sophisticated techniques are used, and on the source of the data, method of collection and quality of the data used, and the impact that the algorithm has on private citizens. Our audit framework intends to assist in making algorithms, and to foster debate on the potential risks of algorithms. Assessors and auditors will be able to use this audit framework in the future to assess algorithms in a consistent and uniform manner.

1.2 Recommendations

In order to ensure that the cabinet has a clear understanding of both the extent to which algorithms are used by central government and how they are used, and in order to provide a clear point of reference, we urge the cabinet to:

- publish clear, consistent definitions of algorithms and quality requirements for algorithms.

In order to ensure that algorithms are used and refined in a responsible way, we urge the cabinet to:

- ensure that the audit framework is translated into practical quality requirements for algorithms;
- ensure that all relevant disciplines are involved in the development of algorithms;
- ensure that clear information is produced now and in the future on the operation of IT general controls[6];
- document agreements on the use of algorithms and make effective arrangements for monitoring compliance with these agreements on an ongoing basis.

We found that private citizens do not play a prominent role in the use of algorithms. We therefore make the following recommendation to the cabinet:

- provide insight in algorithms for citizens and explain where and how they can obtain more information about algorithms.

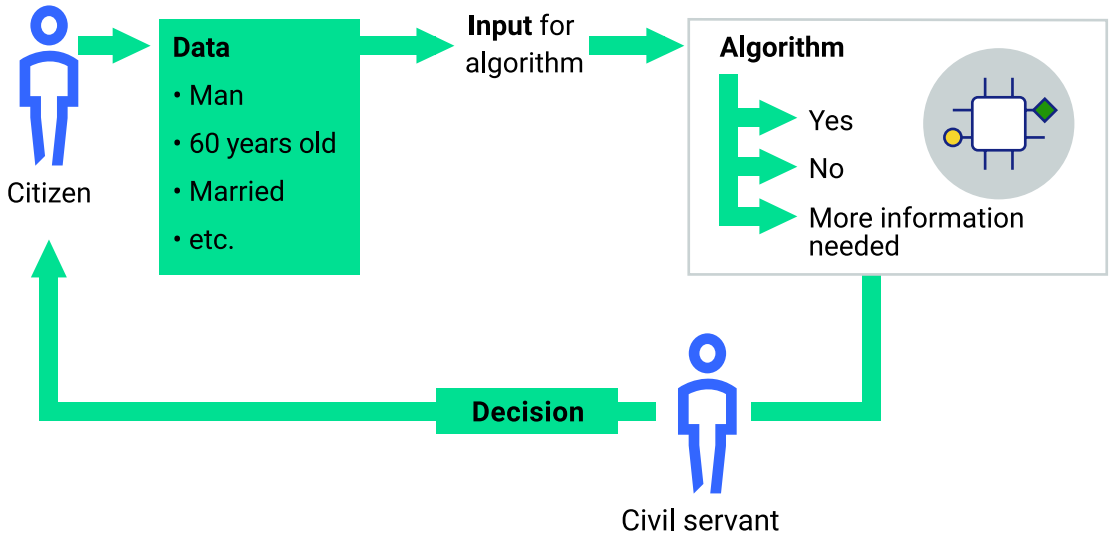
2. About this audit

2.1 Why did we perform this audit?

The central government has been using algorithms for decades. 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.[7] 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.

An algorithm is a set of rules and instructions that a computer follows in order to solve a problem or answer a question

Example: A person applies for benefits. Is he entitled to one?



Algorithms are growing ever more popular, thanks to advancing computerisation and digitisation, algorithms are growing ever more popular. 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 are able to 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.'[8] 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.

In undertaking this audit, 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.

2.1.1 Opportunities for 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.[9] The government is planning to invest €23.5 million in 2021 in the Dutch AI Coalition, a public-private partnership in artificial intelligence. This money has been earmarked for research into artificial intelligence and the development of applications. It is clear from the ministries responses to our audit questions that there is a general agreement in central government about the many new opportunities offered by AI. 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.

2.1.2 Threats posed by algorithms

The use of algorithms by government organisations also poses a number of threats. Four of these are described below.

1. First of all, 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).
2. 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.
3. A third 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.
4. Lastly, 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.

2.1.3 Demystification

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.

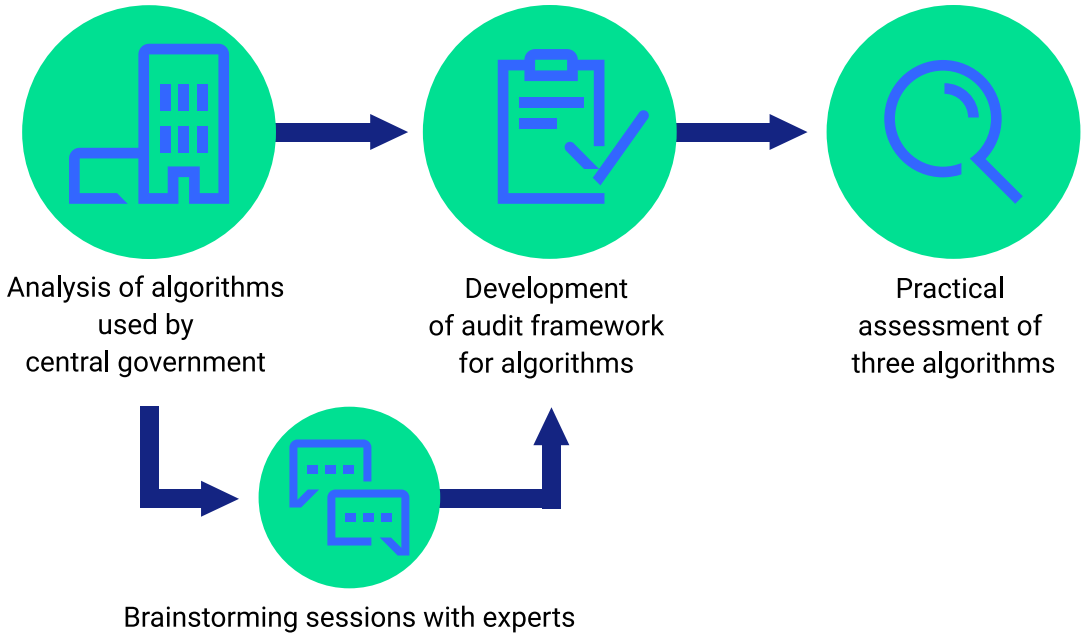
Currently, most algorithms take the form of instructions that a computer follows with the help of data in order to reach a decision. At the same time, they are becoming both more complex and faster-acting. Combined with the potential for social unrest, this development has created a growing need among auditors and regulators for clear guidelines and assessment criteria that they can use to analyse and assess algorithms.

2.2 What did we audit and how was our audit performed?

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 the national government and its associated organisations.

A predictive algorithm is used to answer the question 'What's going to happen next?', whereas a prescriptive algorithm is used to answer the question 'What needs to be done?' 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.[10] Appendix 1 to this report includes a detailed description of our audit methods. We wish to underline that we did not seek to perform a full analysis of all the algorithms used by central government in this audit.

Our audit of algorithms consists of three components



Analysis

Our audit began when we asked the ministries to identify relevant applications of predictive and prescriptive algorithms. We made clear that, for the purpose of this audit, we wished to receive information about algorithms that have both:

- a *predictive* or *prescriptive* function, and
- 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. Our audit aimed to answer the following audit questions.

1. For which activities and processes do central government and its associated organisations use algorithms, what types or categories of algorithms are there, and what are the risks and effects associated with the use of algorithms? (section 3.2)?
2. How do central government and its associated organisations manage the operation and control the quality of algorithms? (section 3.3)?

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.[11] The results of the session are recorded in chapter 4 of this report.

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 or not 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.

2.3 Format of this report

This audit report consists of three parts. Chapter 3 describes the results of our analysis of the use of algorithms in central government and its associated organisations. Chapter 4 describes the development of the audit framework, its five component aspects and the brainstorming session held as part of this audit on 22 September 2020. Chapter 5 lists the main observations and issues that emerged from the practical test of our audit framework. Chapter 6 presents our conclusions and recommendations.

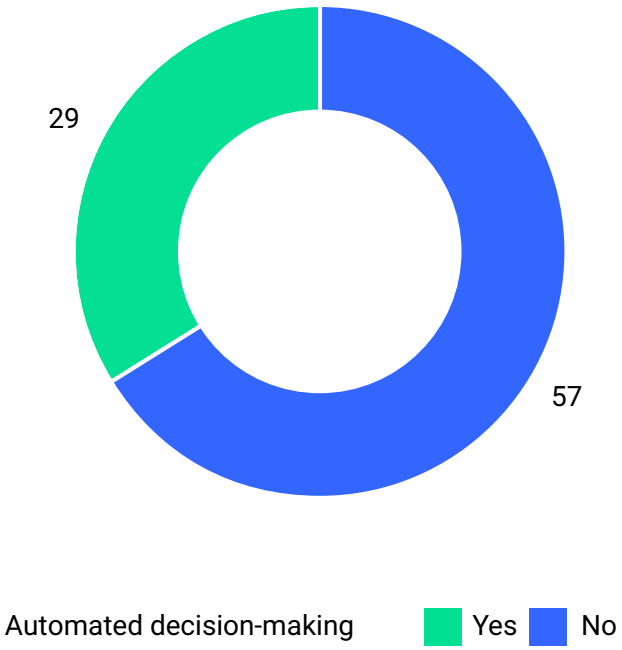
3. Understanding algorithms

3.1 Overall picture 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.

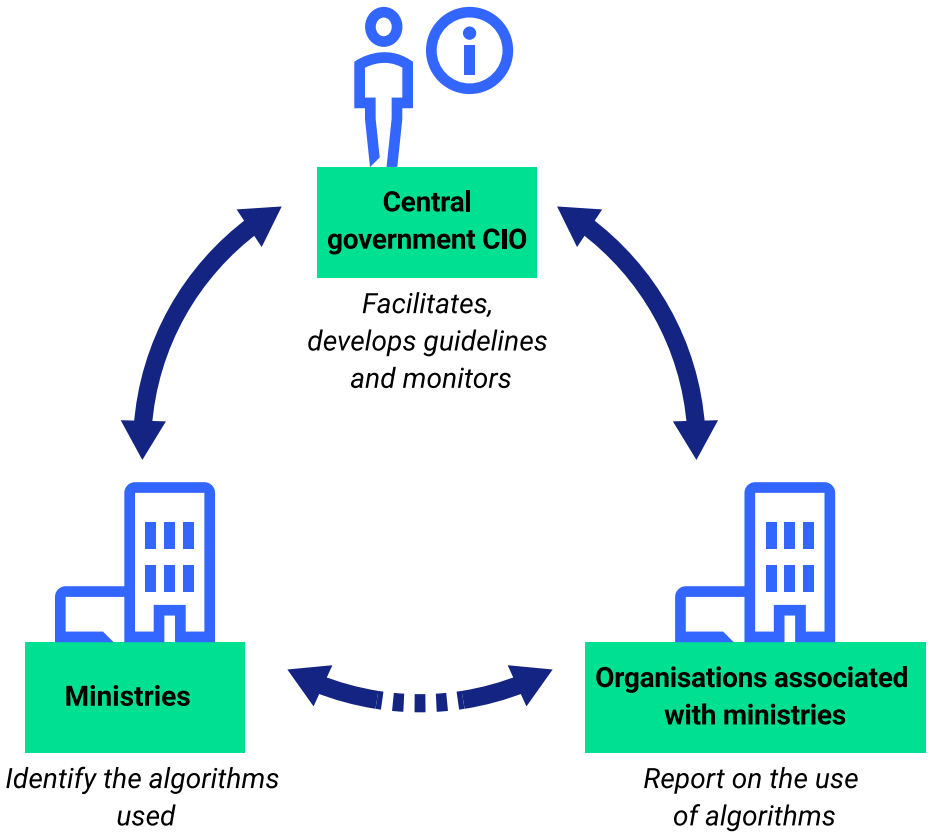
The majority of the algorithms do not use automated decision-making



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. 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.[12] The number of algorithms used is not necessarily a 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.

Ministries and central government CIO do not have a clear picture of the algorithms used



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, 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.

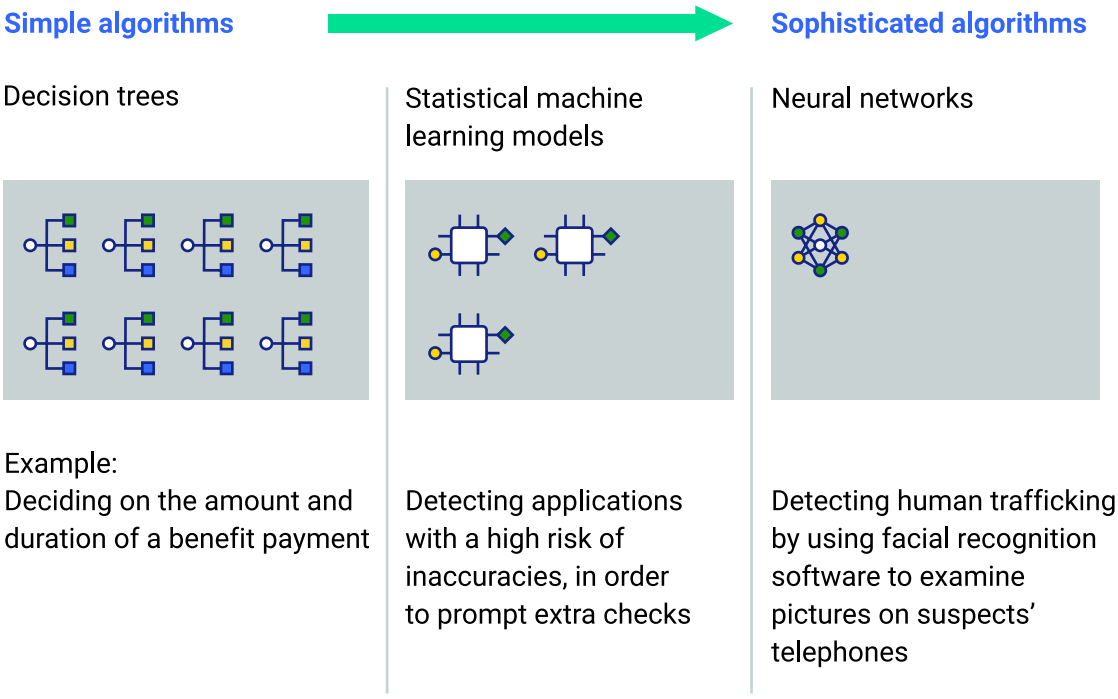
3.2 For which activities and processes do central government and its associated organisations use algorithms, what types or categories of algorithms are there, and what are the risks and effects associated with the use of algorithms?

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.[13] The classification is based on the complexity of the algorithms, ranging from simple to complex. 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,[14] on the other hand, is a complex algorithm. 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.

Central government uses mainly simple algorithms and hardly any sophisticated algorithms



Sitting between these two ends of the scale are algorithms with 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 (see the above figure). 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.

First purpose: automating administrative activities and implementing simple legislation

A part of the algorithms are used to automate routine human activities. The government makes widespread use of such algorithms. This may generate big efficiency gains, in particular 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.

Second purpose: 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. In this way 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 due to the fact that the algorithm only exhibits a prepatory function: it performs an analysis that an expert assesses before taking a final decision.

Third purpose: 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 an official checks the applications. The algorithm informs the official which applications are likely to take more time, without automatically deciding whether or not 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[15] uses these checks a lot, for example with 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, on the basis of their professional judgement, whether to follow this recommendation. In other words, there is no automated decision-making 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.

3.3 How do central government and its associated organisations manage the operation and control of the quality of algorithms?

Our analysis shows that the way in which algorithms are managed by central government is governed by general standards and guidelines, in particular the General Data Protection Regulation (GDPR) and the Government Information Security Baseline. Both ministry officials and the staff of associated organisations would like to see a set of standards or guidelines adopted specifically for algorithms that would reflect the tenor of the wider political and social debate about algorithms.

Officials and staff find it difficult to decide how to manage the operation and quality control of algorithms on a day-to-day basis. Many ministries and associated organisations need an assessment framework in order to gain more control over the use of algorithms, especially because the specific risks attached to algorithms are not always known or clear.

Virtually all the ministry officials who completed our questionnaires said that they would like to see the government draw up a single set of guidelines governing the use and risk management of algorithms. The responsible ministers could set many minds at ease by adopting a common position that meets the needs of both internal and external stakeholders. This could be done by creating a single set of government-wide assessment criteria. In response to our questions, officials of three ministries said that it would not be possible to work with generic standards. When we assessed the three algorithms in practice, we found that the risks were fairly generic. Our audit shows that a generic set of assessment criteria could be used by central government.

4.
An audit framework for algorithms

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. 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.

Virtually all ministries are currently working on standards and guidelines for assessing algorithms. A number of non-governmental organisations are also working on the same issue, among them NOREA, the Dutch professional association of IT auditors, and large accounting firms. The audit framework that we developed 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 forms part of this audit report and is publicly available at: www.rekenkamer.nl/algoritmes-toetsingskader.

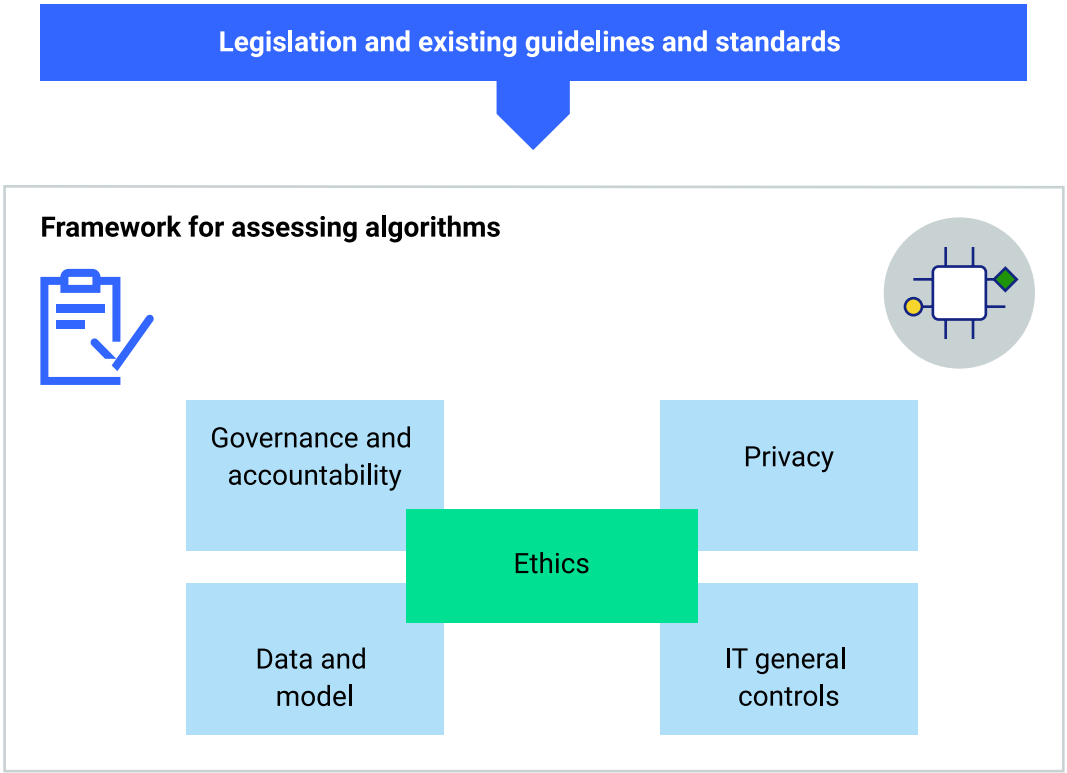
4.1 Five perspectives

Our audit framework contains five different perspectives for investigating algorithms:

1. governance and accountability;
2. model and data;
3. privacy;
4. IT general controls (ITGC);
5. ethics.

Rather than constituting a separate aspect, ethics are interwoven with all the other four aspects. Our audit framework is based on existing standards and guidelines (see Appendix 2). It provides concrete answers to the questions of which aspects need to be assessed, which risks are associated with algorithms, and the audit questions that we wish to answer.

Our audit framework is based on five perspectives



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).[16]

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 minimalisation, and whether or not 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. This may mean that certain aspects do not apply to a specific algorithm.

Privacy

Some algorithms use personal data, including special category personal data.[17] 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.

IT general controls (ITGC)

IT general controls (ITGC) 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.

4.2 Brainstorming session: terminology and definitions

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:

1. data-driven;
2. data quality;
3. artificial intelligence and algorithms;
4. artificial intelligence in central government;
5. transparency.

Appendix 1 contains a report of the brainstorming session.

5. Practical test of three algorithms

5.1 Selection of algorithms

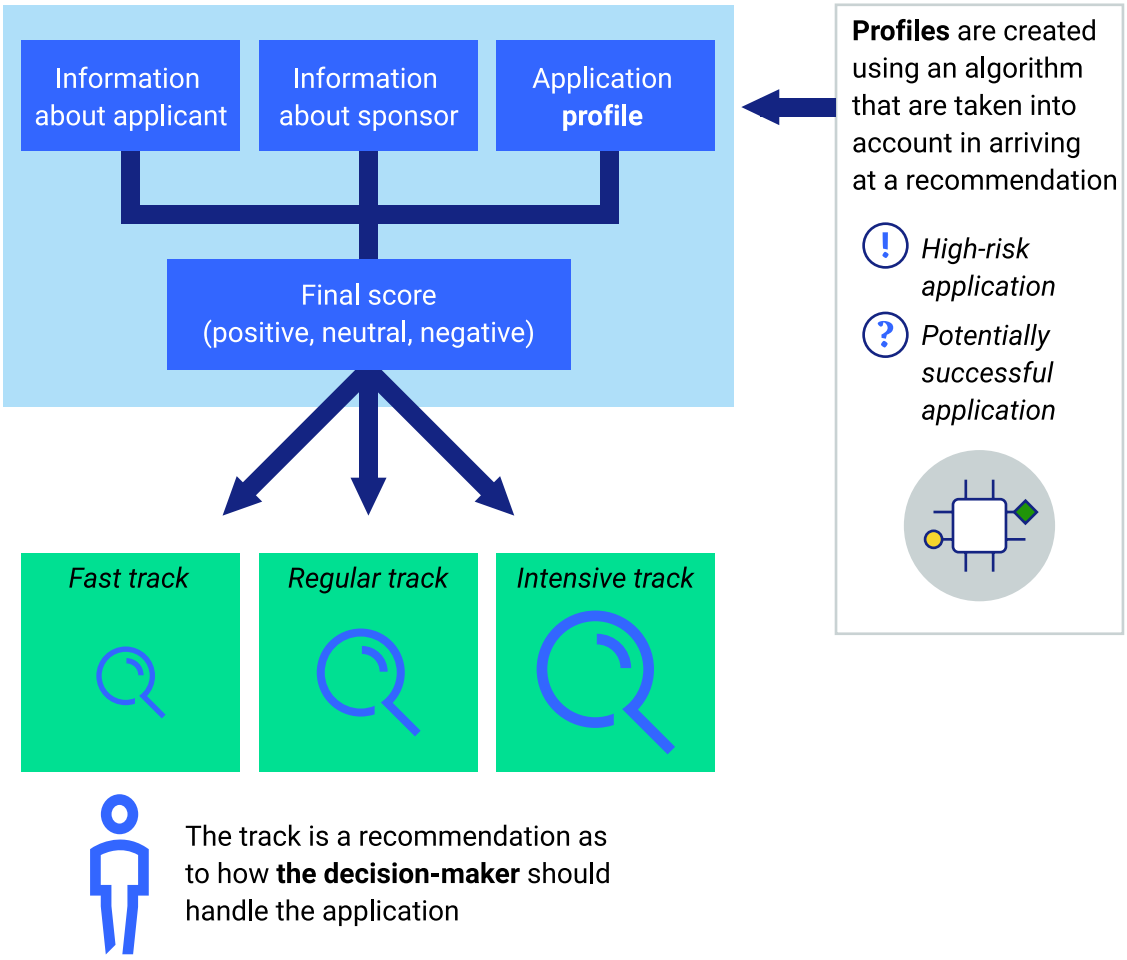
We wanted to submit our audit framework to a practical usability test by assessing three algorithms. We also wanted to improve our framework. We did not seek to arrive at any individual judgements, which is why we generalised our findings. A further objective was to collect more information on the risks 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.

We tested our audit framework by applying it to three specific algorithms:

1. A decision tree designed to make recommendations for checks or extra checks of applications from private citizens;

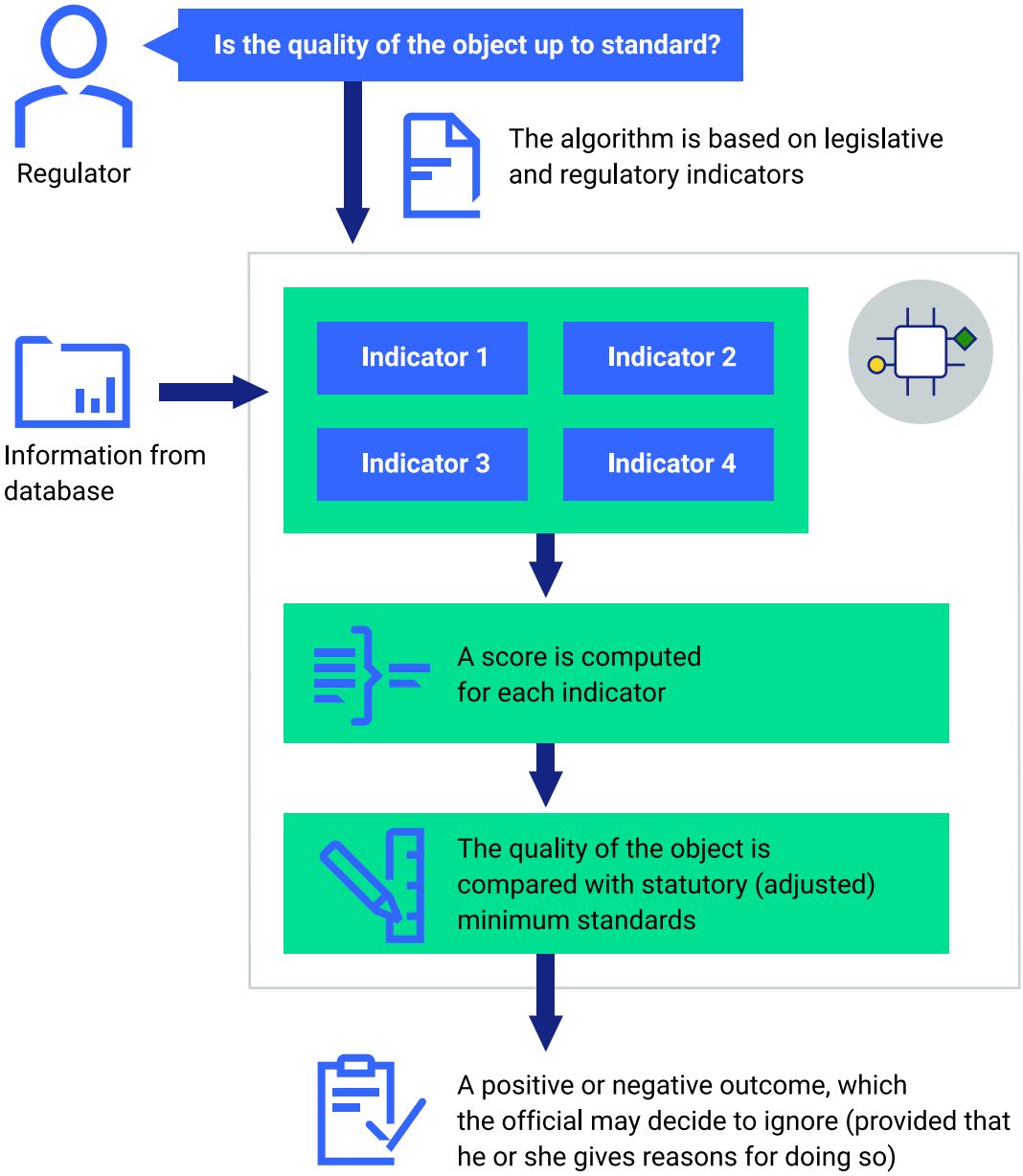
An algorithm based on simple decision-making rules can help officials to check applications more efficiently

A person or company applies for a grant, a travel document or a benefit. Does this application need checking or extra checking?



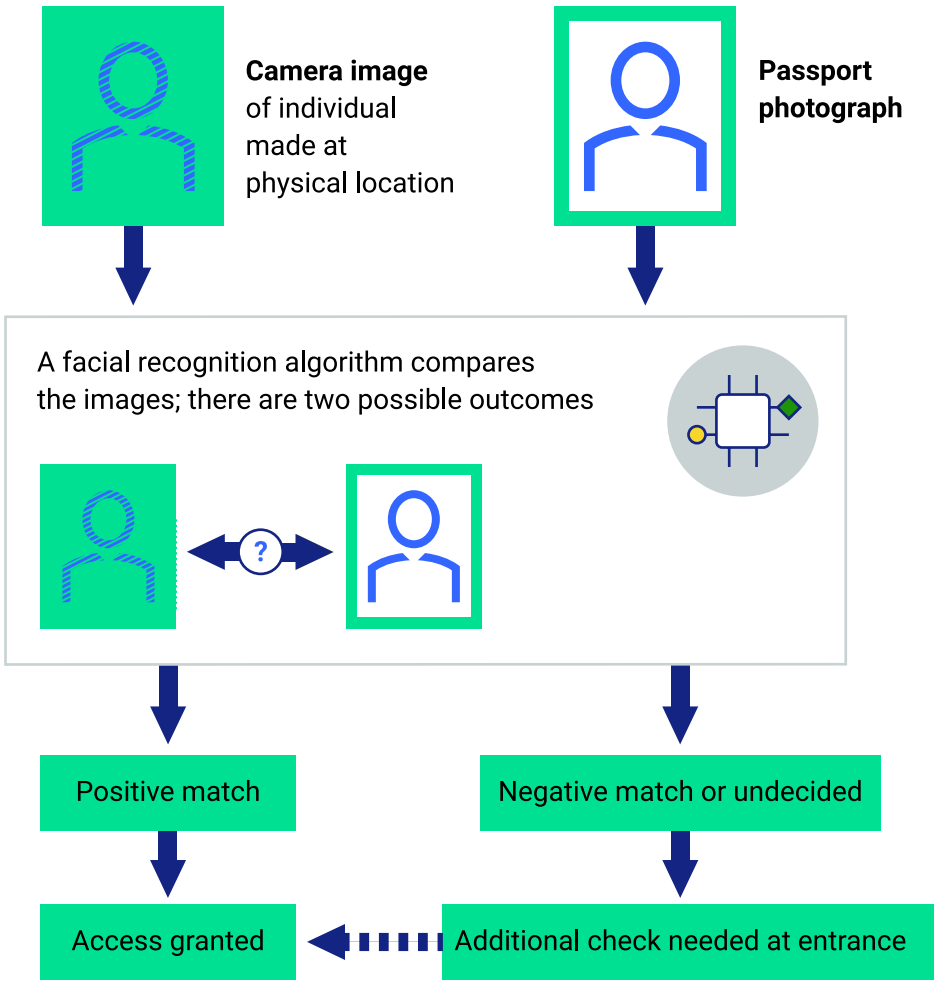
2. An assessment system for detecting non-standard objects, generating information for regulators and inspectors;

Algorithms assist regulators and inspectors in making assessments



3. A facial recognition system for granting individuals physical access to a site or building.

Algorithms can use facial recognition to help decide whether to grant individuals physical access



We selected these three algorithms for the following reasons:

1. they are predictive and/or prescriptive algorithms used on a day-to-day basis;
2. they have a substantial impact on private citizens and businesses;
3. they use different techniques.

5.2 Main observations

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.[18]

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 system of life cycle management has been adopted for algorithms.[19] 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.

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.

1. 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.
2. 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.

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.

In the cases of the algorithms we assessed, 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. They did this 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')[20] and *MijnOverheid* ('My Government')[21] 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.

IT General Controls (ITGC)

It is clear from the limited amount of documentation that we received from the auditees that, of the four perspectives of our audit framework, the ITGC requirements 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,[22] 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.

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.

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 four aspects. We identified four themes from an ethical perspective, based on existing sources (see Appendix 2) and standards:

1. respect for human autonomy;
2. the prevention of harm;
3. fairness (a fair algorithm);
4. explainability and transparency.

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 supports officials.

The prevention of harm

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.

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.

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 conclusion

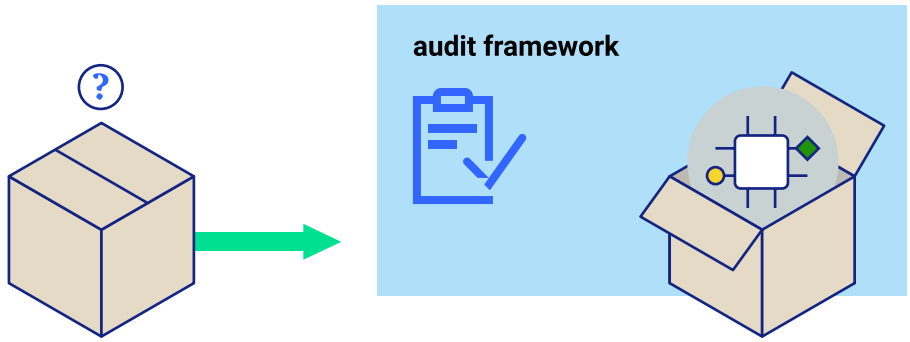
In order to assess whether an algorithm adheres to the ethical principles of fairness, explainability and transparency, independent assessors must be able to identify and check the data used. In the case of one algorithm, the data needed to comply with privacy legislation was not stored. This means that,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

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.

An algorithm is not a black box



Our main conclusion based on the algorithms we analysed 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.

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 an other choice to make. 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)[23] and COBIT 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 only 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.

6.2.1 First recommendation: 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:

- for automating administrative work and simple legislation;
- for facilitating and improving operational management and/or service delivery;
- 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 Private citizens are insufficiently taken into account

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 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. 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.

6.4.1 Second recommendation: 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

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.

6.5.1 Third recommendation: 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.

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.

6.5.2 Fourth recommendation: 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.

6.5.3 Fifth recommendation: 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.

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 (www.avgregisterrijksoverheid.nl) 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 section 6.4.1.

IT General Controls (ITGC)

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.

6.5.4 Sixth recommendation: 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. They can do this 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.

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.
State Secretary's response and Court afterword

The State Secretary for the Interior and Kingdom Relations responded to our report on 22 December 2020, writing in his capacity as the person responsible for coordinating IT matters in central government, and also on behalf of his colleagues.

7.1 Response of the State Secretary for the Interior and Kingdom Relations

In his response, the State Secretary for the Interior and Kingdom Relations (the State Secretary) states that he accepts and values our conclusions. He describes our recommendations as constructive. Below is a summary of his response to our recommendations. The full text of his response (in Dutch) can be found at www.rekenkamer.nl. This chapter concludes with our afterword.

Response to recommendations

Your recommendations will help to improve the delivery of services to the people for whom the government works and the relevant operational processes.

1. "Publish clear, consistent definitions and quality requirements."

We are seeking to define a consistent, common set of terms and specific quality requirements for algorithms with the aid of knowledge pooling and the meetings planned in line with the Dutch Digitalisation Strategy, among other initiatives. Acting in conjunction with the Minister for Legal Protection and the State Secretary for Economic Affairs and Climate Policy, we have performed an exploratory analysis

that looked at issues such as avoiding fragmentation, standardising monitoring procedures, and making use of both public-sector and private-sector expertise. The Dutch House of Representatives was informed of the results at the time of your audit,[24] and discussed these with a number of ministers and state secretaries.

In these efforts, it is important to strike the right balance between the added value achieved from government-wide uniformity on the one hand and the need to take account of the specific requirements of individual ministries and executive agencies on the other.

2. "Inform private citizens about algorithms and explain how they can obtain further information about algorithms."
The existing guidelines for the use of algorithms by governments are currently being refined and evaluated. In addition, a model is being developed for assessing the impact of algorithms on human rights. Both national and European legislation contain information about predictive or prescriptive algorithms.

Your report cites SyRi as an example. Page 5 of the report states that the SyRi system was used by the government (notably by the Employee Insurance Agency and the Tax and Customs Administration) as a means of detecting fraud with the help of algorithms. As this sentence creates an impression that SyRi is in widespread and generic use for performing routine checks, we want to explain the context and use of SyRi. SyRi is a system that compares data files of different government organisations (both central and local) based on the Work and Income (Implementation Organisation Structure) Act. The system was used for a small number of specific joint projects for preventing and reducing tax and social security fraud, infringements of labour laws and related instances of abuse of the law. On 5 February 2020, the court ruled that the use of SyRI represented an unacceptable infringement of citizens' privacy rights. Following the court's ruling, the government immediately stopped using SyRI.

3. "Document agreements on the use of algorithms and make effective arrangements for monitoring compliance with these agreements on an ongoing basis."
Detailed agreements on the use and monitoring of algorithms have been reached in close cooperation between various ministries. Tangible results include the Strategic Action Plan for Artificial Intelligence, a policy letter on public values, and safeguards against the risks posed by data analysis performed by government.

4. "Ensure that the audit framework is translated into practical quality requirements for algorithms."

Working in conjunction with the Netherlands Court of Audit and the Central Government Audit Service, the government is seeking to translate the audit framework into practical quality requirements for algorithms. Where artificial intelligence (AI) algorithms are concerned, the reliability and quality of data must also be included in the assessment, as AI algorithms use data. The government is seeking to put the necessary safeguards in place by setting an agenda and establishing a project team to ensure that the results of the exploratory analysis and the Court of Audit's audit report are taken into account during both the development and implementation stages, and to agree on a broadly supported agenda for action on the standardisation and monitoring of algorithms. We will need to perform further research into the potential effects this could have in terms of the administrative burden and the practicality demands placed on both local and national government officials, and into the amount of time required for implementation, taking due account of existing mechanisms and the autonomy of the ministries and executive agencies.

5. "Ensure that all relevant disciplines are involved in the development of algorithms."

Interdisciplinary cooperation is standard procedure in the development of policy, legislation, operating processes and the resultant algorithms and the monitoring of such algorithms, in accordance with the government's comprehensive decision-making framework. The use of policy tools, algorithms and other instruments is based on the law and is confined by the limits of the agreed framework. At the same time, the ideal mix of disciplines for development and other processes inevitably differs from one individual instance to another, depending on the availability of staff capacity, resources and time.

6. "Ensure that clear information is produced now and in the future on the operation of IT general controls."

IT general controls are crucial to the operation of both conventional systems and algorithms. Additional reporting mechanisms or audit reports can help the responsible authorities to understand and monitor such systems. Activities undertaken in parallel with these can also help CIOs and their staff to track the operation of the systems in question. At the same time, it is important to remain aware of the dynamic nature of the operating environment and hence to strike a balance between controls on the one hand and the organisational or administrative workload they entail on the other.

You performed a practical test on three specific algorithms using your audit framework. You then generalised on the basis of your findings. I would like to point out that this means that your audit report pays less attention to high-quality applications of algorithms by central government in which ethical standards form one of the organisational principles.

While acknowledging the importance of safeguarding the rights and freedoms of private citizens, we will need to acquire more information – partly in order to bolster the regulatory system – about the scope available to ministries and executive agencies to experiment with algorithms, so that both regulators and those involved on the operational side can learn from each other's experiences.

As you write in your report, collaboration has recently been strengthened, thanks in part to the brainstorming session held on 22 September 2020, which was attended by representatives from a number of ministries, external experts, and staff from the Netherlands Court of Audit and the Central Government Audit Service.

Thank you for performing this audit and for helping in this way to foster a greater understanding of algorithms. Your findings will help to improve both the delivery of public services and the implementation of government policies."

7.2 Court afterword

We would like to thank the State Secretary for his response.

In responding to our recommendation to publish clear, consistent definitions and quality requirements, the State Secretary stresses the importance of striking the right balance between the added value achieved from government-wide uniformity on the one hand and the need to take account of the specific requirements of individual ministries and executive agencies on the other.

Our audit team found that, in virtually all ministries, the officials responsible for the development and application of algorithms had a considerable need for greater uniformity in the terminology and quality guidelines and standards used. Officials at the Ministry of the Interior and Kingdom Relations (specifically, the staff of the central government CIO) could play an important role in this. Uniformity helps both in the sharing of knowledge and in achieving consistency in quality, and in doing so can

create the space required to find customised solutions for individual ministries and executive agencies. It can also help to boost collaboration between ministries in making arrangements about the way in which algorithms are used and monitored (see the third recommendation). It is absolutely vital that such arrangements are consistent, verifiable and binding.

We would like to point out that our audit framework covers risks applying to all sorts of algorithms, irrespective of the context in which they are used by the ministry or executive agency in question. Our assumption is that our audit findings can be used to help all ministries use algorithms in a responsible manner.

The impact assessment and legislation referred to by the State Secretary (in connection with the second recommendation) undoubtedly form a good starting point. In urging the government to inform private citizens better about the use of algorithms, we also feel that sufficient attention needs to be paid (and at an early stage) to the practical aspects, and that the general public should be actively informed about the possibilities in this connection.

We welcome the State Secretary's interest in translating the audit framework into practical quality requirements for algorithms (see the fourth recommendation). Data reliability and quality are highly critical issues in relation to all types of algorithms. It is true that AI algorithms are particularly important in this respect, not only because of the huge volumes of data they use, but also because it is not always possible to keep track of how the data is processed and how algorithms reach their conclusions.

We also welcome the State Secretary's confirmation that interdisciplinary cooperation is standard procedure in policy development, and hence also in the life cycles of algorithms (see the fifth recommendation). However, our auditors found that this had not yet been translated to a sufficient degree into day-to-day practice. While readily accepting that financial or practical considerations may play a role, we wish to stress the advantages to be gained from an interdisciplinary approach in terms of the mitigation of risks, particularly in the relatively long term.

While all sorts of audit reports and statements can provide greater information about IT general controls (see the sixth recommendation), what we believe to be critical is the level and quality of access and data security, the life cycle management and the continuity mechanisms, both in general and in relation to algorithms, all of which need to be safeguarded by the owner or manager.

The State Secretary acknowledges the importance of safeguarding the rights and freedoms of private citizens in relation to the use of algorithms. He rightly links this with a desire to know more about the scope available to ministries and executive agencies to experiment with algorithms – a desire that may give rise to certain dilemmas. While we are sympathetic to this line of reasoning, we believe that neither of the two considerations should be allowed to impinge on the other. Moreover, the amount of information published and the channels through which such information is made available, including information on the nature and scope of experimental algorithms, are both vitally important. In line with the point made about quality requirements, we wish to stress that our audit framework covers risks applying to all sorts of algorithms, at each stage of development or use.

In performing this audit, we hope to help dispel certain understandable concerns among the general public about the government's potential use of unverifiable, decision-making algorithms. We found that the government currently only makes very limited use of decision-making algorithms and that the algorithms we assessed were indeed verifiable. This is not to say, however, that there is no reason for any concern. Things are moving at a rapid pace and we did not check whether the lists of algorithms supplied to us by the ministries were complete. Nevertheless, not only are algorithms made by human action, they are also open to verification by human action concerning of their impact on private citizens. And this is how it should be.

We will keep track of the progress made in implementing our recommendations and will continue to focus on algorithms as one of our audit topics in the years ahead.

Appendices

Appendix 1
Audit methods

Understanding algorithms

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, we analysed the types of algorithms used by central government and the activities for which they are used. 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 10 algorithms, but this was merely an indicative number.

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.[25] It classifies algorithms on the basis of the following characteristics, among others:

- complexity (low-high);
- technical transparency (low-high).

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, in accordance with the following scale:

- descriptive;
- diagnostic;
- predictive;
- prescriptive.

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.

Audit framework and practical test of algorithms

We developed an audit framework based on existing standards and guidelines, and the relevant literature (see Appendix 2). The audit framework was then refined based on the outcome of a brainstorming session and a practical test (in which we assessed three specific algorithms).

Brainstorming session

There are numerous guidelines and standards governing individual features of algorithms. For example, the GDPR for privacy and the Government information security baseline for information security. There is no single comprehensive audit framework covering all aspects of algorithms, however. Moreover, there is no common terminology for discussing algorithms. The officials taking part in our analysis said that they would like to see more uniformity in the definitions and terminology used. What exactly is an algorithm? What does explainability mean? And for whom does an algorithm need to be explainable? What do we mean by transparency?

In order to meet this need, 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. The objective of the session was to achieve greater uniformity in the terminology used for algorithms. Greater uniformity

would also help in creating guidelines for the day-to-day use of algorithms, and in arranging a suitable form of accountability. The idea is that discussing the characteristics and definitions of algorithms from different viewpoints (i.e. legal, technical, policy, scientific and regulatory), and learning from each other's experiences should produce a clearer picture and a better understanding of algorithms. Thirty experts from both within and beyond central government took part in the session. A report of the brainstorming session follows below.

Report of the brainstorming session

Objective

There are numerous guidelines and standards governing individual features of algorithms. For example, the GDPR for privacy and the Government information security baseline for information security. There is no single comprehensive audit framework covering all aspects of algorithms, however. Moreover, there is no common terminology for discussing algorithms. What exactly do we mean by algorithms? *What is explainability, and what does transparency mean? What's the difference between them, and for whom should an algorithm be explainable? And what does bias mean? Wasn't bias always there? And will things become more complex now that the bias is in the algorithm rather than in the human brain?*

In order to achieve greater uniformity in the terminology used for algorithms, we organised a brainstorming session on algorithmic data analyses on 22 September 2020 in conjunction with 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.

Review & results

Over thirty experts from both within and beyond government took part in the brainstorming session. Based on their different roles, backgrounds and expertise (i.e. legal, technical, policy, scientific and regulatory), they discussed five themes: data-driven work practices; data quality; artificial intelligence and algorithms; artificial intelligence in government, and transparency. What were some of the noteworthy points that emerged?

- Although algorithms have been used for many years both within and beyond central government, they are regarded as 'scary' due to the negative press they have received. The technology used for algorithms should be made more comprehensible, so that people can gain a better understanding of how they are used, thus demystifying them;

- The need for a uniform audit framework or uniform guidelines springs from a desire to create a common set of more detailed foundations in central government.
- Both context and purpose must be taken into account in order to create a better understanding of algorithms.
- Many of the above themes are too broad and too abstract. Reaching agreement on a single, uniform definition is not a realistic aim. Agreement could be reached, however, by splitting up the definition into its main constituent parts and spelling these out in detail.
- Algorithms and data require government-wide management. This is because there is a growing tendency for government to operate through networks of government organisations.
- In those cases where the management of algorithms or parts of algorithms is outsourced, government organisations do not always closely monitor their operation, and they have less control over some of these algorithms than over those managed in-house.
- Even if an algorithm is perfect 'on paper', it is humans who make it and who decide on the data used by the algorithm. There is never a watertight guarantee that models (or humans) have no bias and therefore do not discriminate. Politicians need to take this into account in order to adopt effective, practicable controls.
- Central government should set certain minimum requirements for the use of algorithms, so that they are used in a responsible way.

Appendix 2
Reference list and sources used for audit framework

This reference list includes a selection of the main sources used, but is not complete, due to the large number of publications available on this subject.

Parliamentary documents

- Ministry of Economic Affairs and Climate Policy (2019), *Kamerbrief van Minister van EZK over Strategische Actieplan voor Artificiële Intelligentie*, Dutch House of Representatives, 8 October 2019, Parliamentary Paper 26 643, no. 640
- Government of the Netherlands (2019), *Strategic Action Plan for Artificial Intelligence*, 8 October 2019
- Ministry of Justice and Security (2019), *Kamerbrief van Minister van J&V over Waarborgen tegen risico's van data-analyses door de overheid*, 8 October 2019, Parliamentary Paper 26 643, no. 641
- Ministry of the Interior and Kingdom Relations (2019), *Kamerbrief van Minister van BZK over AI, publieke waarden en mensenrechten*, Dutch House of Representatives, 8 October 2019, Parliamentary Paper 26 643, no. 642
- Government of the Netherlands (2020), *Kabinetsreactie op het onderzoek 'Toezicht op het gebruik van algoritmen door de overheid'* Date 20 April 2020, appendix to Parliamentary Paper 35 212, no. 3

National

- Central Government Audit Service (2018), GITC framework based on Civil service baseline information security 2017
- *BIR 2017, Civil service baseline information security,* fully aligned with international ISO/IEC 27002 standard
- BIO, *Government Information security baseline,* effective as of 1 January 2019 (published in the Government Gazette on 23 May 2019), replaces Civil service baseline information security (BIR) 2017
- Amsterdam local authority, *Modelbepalingen voor gemeenten voor verantwoord gebruik van Algoritmische toepassingen*, https://www.amsterdam.nl/wonen-leefomgeving/innovatie/de-digitale-stad/grip-op-algorithmes/
- Hooghiemstra & Partners (2019), *Onderzoek Toezicht op het gebruik van algoritmen door de overheid,* Hooghiemstra & Partners
- Waag (2020), *Algoritme: de mens in de machine*, Waag
- Frans van Bruggen and Joep Beckers (2020), *Nut en noodzaak van toezicht op artificiële intelligentie*, Tijdschrift voor Toezicht

- Montaigne Centrum voor Rechtsstaat en Rechtspleging, Utrecht University (2020), Juridische aspecten van algoritmen die besluiten nemen, Een verkennend onderzoek, Montaigne Centrum
- Dialogic (2020), *Gebruik van en toezicht op AI-toepassingen in telecom-infrastructuren, Advies aan de toezichthouder over inrichting van risico gebaseerd AI-toezicht,* Radiocommunications Agency Netherlands

EU & International

- High-level expert group on artificial intelligence set up by European Commission (2019), *Ethics guidelines for trustworthy AI*, European Commission
- Michael Veale (2019), *A Critical Take on the Policy Recommendations of the EU High-Level Expert Group on Artificial Intelligence*, Faculty of Laws, University College London and the Alan Turing Institute
- National Audit Office UK (2016), *Framework to review models*
- Anna Jobin, Marcello Ienca and Effy Vayena (2019), *Artificial Intelligence: the global landscape of ethics guidelines,* Health Ethics & Policy Lab, ETH Zurich
- Thilo Hagendorff (2020), *The Ethics of AI Ethics: An Evaluation of Guidelines, Minds and Machines*
- European Commission (2020), *Whitepaper on Artificial Intelligence - A European approach to excellence and trust*, European Commission
- Daten Ethik Kommission (2019), *Opinion of the Data Ethics Commission, Daten Ethik Kommission*
- OECD (2019), *Recommendation of the Council on Artificial Intelligence,* OECD
- Geron, A. (2017), *Hands-On Machine Learning with Scikit-Learn and TensorFlow*
- Hastie, T., Tibshirani R, and Friedman, F. (2009), *The Elements of Statistical Learning*
- Thomas L.C., Oliver R.W., and Hand D.J. (2005), *A survey of the issues in consumer credit modelling research*, Journal of the Operational Research Society, 56, 1006-1015
- ISACA (2018), *Auditing Artificial Intelligence*, ISACA
- ISACA (2012), COBIT 5, *A Business Framework for the Governance and Management of Enterprise IT*, ISACA
- ISACA (2012), *COBIT 5, Enabling Processes*, ISACA

Online sources

Kennisbank openbaar bestuur, *Artificiële Intelligentie en publieke waarden,* https://kennisopenbaarbestuur.nl/thema/artifici%C3%ABle-intelligentie-en-publieke-waarden

Appendix 3
Audit framework for algorithms

The audit framework that we developed as part of our *Understanding Algorithms* audit is a practical tool for managing the main risks posed to central government by the use of algorithms. We made use of existing standards, guidelines and legislation. The framework consists of five different perspectives:

1. governance and accountability;
2. model and data;
3. privacy;
4. IT general controls (ITGC);
5. ethics.

We identified the main risks relating to each of the above perspectives, and linked the elements of the assessment and our audit questions to these risks. Once all the various questions have been answered and scores allotted, the result is a picture of the extent to which risks relating to a particular algorithm have been mitigated. The degree of risk associated with a specific algorithm depends on the sophistication of the algorithm and its impact on private citizens.

Before the framework can be used, a number of general questions must first be answered. The information provided in answering these questions generates a general impression of the algorithm and its context. It is this context and general impression that determine which questions are selected from the audit framework for the purpose of assessing the algorithm in question.

General questions

1. What is the name of the algorithm or the system of which the algorithm is part?
2. In which operating process for which product or service is the algorithm used?
3. Does the algorithm make use of personal data (GDPR)?
4. Is it a learning algorithm, i.e. an algorithm that evolves and improves over time by using data and/or experiences?
5. Does the algorithm advise or support human activities or decisions, or does it act autonomously or automatically without any human interference?
6. What technology and/or what application or software does the algorithm use?
7. Which data and data sources does the algorithm use?

The audit framework

Governance and accountability		
Risk	**Audit question**	**Ethical principle[26]**
There can be no management or accountability without clarity about the purpose of an algorithm.	Does the algorithm have a clearly defined purpose?	4.2
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.	Are regular documented assessments made (at the start of a business case) about the management of the risks associated with the use of the algorithm?	4.1
There is a greater risk of error without adequate resources in both qualitative and quantitative terms.	Does the organisation have access to sufficient expertise in both qualitative and quantitative terms?	
No full picture of the life cycle, making the algorithm impossible to manage.	Has the entire process (life cycle) surrounding the algorithm been documented?	
Lack of clarity about roles, tasks, responsibilities and powers creates risks.	Have roles, tasks, responsibilities and powers (including ownership) been defined and have these been assigned in practice?	4.1
Performance and quality targets cannot be measured if there is no policy in place.	Is there an agreed and documented policy on quality and performance targets for algorithms?	4.2
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.	Where certain elements or activities relating to the algorithm have been outsourced, have the arrangements made with external suppliers been documented?	4.1
The algorithm cannot be managed without any monitoring, leading to a higher level of risk.	Is the algorithm monitored at regular intervals (at least in relation to availability, performance/quality, safety, compliance with current legislation and regulations, and outsourcing)?	

Model & Data		
Risk	**Audit question**	**Ethical principle**
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.	Does the algorithm have a purpose and has this been translated into practical features with respect to the model and data used? Which particular task or aspect of operational management is the algorithm intended to support?	4.2
Without agreement on the objectives, there is a greater risk of error and differences of interpretation.	Does the algorithm has an agreed purpose, and is this clear and explainable to the owner, developer and user?	4.2
The operation of the algorithm cannot be explained or is difficult to explain.	Is the algorithm explainable and has an attempt been made to strike a balance between the models' explainability and performance?	4.2
The reasons underlying the choices made in the design and implementation of the algorithm can no longer be traced (explained).	Has a record been made of the reasons underlying the choices made in the design and implementation of the algorithm?	4.1, 2.1
No continuity in the process or the performance of activities, due to lack of documentation.	Is there any documentation describing the design and implementation of the algorithm?	4.1
Hyper-parameters were selected at random, and the wrong choices were made in doing so.	Was the selection of hyper-parameters supported by arguments and evidence?	
A lack of transparency for private citizens, businesses and stakeholders; non-compliance with transparency legislation.	Has the model (i.e. the code and mode of operation) been published and is it available to stakeholders? Does the same apply, where possible, to the data used or a description of the data used?	
The algorithm uses automated decision-making even though this is not permitted; or no opportunities for human intervention.	If the algorithm leads to automated decision-making, does it comply with the relevant legislation?	1.1, 2.1
Very limited sources of input mean a higher risk of error and non-compliance with objectives and legislation.	Were the various stakeholders and 'end users' of the algorithm involved in the development process?	3.1
The algorithm does not operate as planned.	Which input-output checks have been performed to safeguard the accuracy and completeness of data processing?	2.1

Model & Data		
Risk	**Audit question**	**Ethical principle**
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.	Is the model updated at regular intervals to bring it into line with current legislation?	
Incorrect training or testing may lead to overfitting or underfitting, or bias.	Have safeguards been put in place regarding the quality of the choices made in relation to training and test data?	4.1
The model leads to undesirable systematic variance for certain individuals, groups or other units (i.e. bias).	Have safeguards been put in place to avoid any bias resulting from the choices made in relation to the model?	3.1, 3.2
There is an undesirable systematic variance (bias) in the data.	Is there no undesirable bias in the data?	3.1, 3.2
A lack of separate processing leads to overfitting, which means that the model cannot be used for new observations.	Have training, test and validation data been processed separately?	
The data is not representative.	Is the data used representative for the application for which the algorithm is used?	2.1, 3.1, 4.1
Dependency on third parties with respect to data used.	Does the government have full control ('ownership') of the data used for the model?	
Violation of basic premises and rules pertaining to data minimalisation and proportionality.	Is there evidence of data minimalisation. Have proportionality and subsidiarity been taken into account?	2.1
The performance metrics are not consistent with the purpose of the algorithm.	Has the quality of the model been documented?	4.2
The data on which the model is based is available only after the outcome has been identified.	Is there evidence of target leakage? That is to say, do the features of the model include the outcome that the model is designed to predict?	
The prediction meets the requisite standard.	Have performance indicators or performance metrics been used?	2.1, 4.2
The model does not always work in practice.	Is the model's output monitored?	2.1

Model & Data		
Risk	**Audit question**	**Ethical principle**
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.	Has the operation of the model or algorithm, including its limitations (i.e. what it can and cannot do) been communicated to external parties?	4.2
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.	Have arrangements been made for the maintenance and management of the algorithm?	

Privacy		
Risk	**Audit question**	**Ethical principle**
Not compliant with statutory regulations under the GDPR.	Is the use of personal data recorded in a register?	2.2
The design of the algorithm does not take sufficient account of the need to protect privacy.	Is there evidence of "data protection by design"?	2.2
Not compliant with statutory regulations under the GDPR.	Has a Data Protection Impact Assessment been performed (if applicable)?	2.2
The algorithm uses automated decision-making even though this is not permitted under the GDPR.	Is there evidence of automated decision-making, and if so, is this permitted?	2.2
Not compliant with statutory regulations under the GDPR; not serving mankind.	Can those involved opt out of automated decision-making (if applicable)?	2.2
Disproportionate use or collection of personal data.	Is there evidence of data minimalisation?	2.2
Unlawful action.	Is data processed in order to discharge a statutory duty?	2.2
Not compliant with GDPR or not fit for purpose.	Is the use of the algorithm to process (special-category) personal data consistent with its original purpose?	2.2

Privacy		
Risk	**Audit question**	**Ethical principle**
Not compliant with statutory regulations under the GDPR.	Have the controller and the data processor for the algorithm and the data used been designated?	2.2
Violation of Article 1 of the Constitution or Article 14 of the ECHR.	Do the data used and the model not lead to discrimination?	2.2
Profiling as defined in Article 4 (4) of the GDPR; risk of contravening the GDPR.	Has an assessment been made of whether there is evidence of profiling and whether this is permitted?	2.2
Not compliant with statutory regulations under the GDPR.	Have arrangements been made for informing, either pro-actively or on request, individuals whose data is processed or used (in relation to both data and algorithm)?	2.2
Not compliant with statutory regulations under the GDPR.	Is the logic behind the algorithm and the data used sufficiently clear to data subjects?	2.2
Not compliant with statutory regulations under the GDPR.	Is the impact of the use of the algorithm clear to data subjects?	2.2
Data subjects are not informed of their rights or of the algorithms and data used.	Is there a publicly available privacy policy describing the data and algorithms used?	2.2

ITGC		
Risk	**Audit question**	**Ethical principle**
Without any logging information, there is no audit trail for tracing when adjustments were made.	Is logging information about the operation of the algorithm recorded and stored in an assessible manner?	
Access rights are no longer up-to-date.	Are there any checks made of whether access rights are up-to-date with respect to the algorithm's operating environment?	2.2
Unlawful access to the algorithm.	Are access rights updated when a member of staff leaves or moves to a different post?	2.2

ITGC		
Risk	**Audit question**	**Ethical principle**
Access rights are issued by unauthorised staff.	Are access rights issued by staff who are authorised to do so?	2.2
Risk of the algorithm being manipulated in cases where access rights are incompatible.	Is there a mechanism for preventing individuals who are entitled to access the algorithm from playing a number of different roles at the same time (segregation of duties)?	2.2
The more users are granted special powers, the greater the risk of manipulation.	Are management accounts generic? Is there a logical relationship between the number of management accounts and the number of managers?	2.2
User groups are difficult to identify.	Are naming conventions used when granting access rights to different user groups or roles? Is this done on a systematic basis?	2.2
Managers and users are difficult to identify.	Are naming conventions used for users and managers, so that they can be identified?	2.2
Unclear who made changes to or worked on the algorithm.	Do managers perform management and ordinary user activities under two different user names?	2.2
The database is open to manipulation if holders of user accounts have access to underlying components.	Do user accounts have access to underlying components?	2.2
The database is open to manipulation if holders of user accounts have access to underlying components.	Is there a strict separation of activities as far as applying for, authorising and processing changes in user accounts and access rights are concerned?	2.2
The database is open to manipulation if holders of user accounts have access to underlying components.	Are passwords managed interactively, and are they of adequate quality?	2.2
Unauthorised access, changes, damage to and/or loss of data. Non- compliance with the law.	Are changes made to the code of the algorithm verifiable? (for example, are changes tested and approved or authorised?)	2.2
Unauthorised access, posing a risk of the algorithm being manipulated (changes, damage, loss of data).	Is the algorithm protected, so that there is no risk of unauthorised access, changes, damage and/or loss of data?	2.2

ITGC		
Risk	**Audit question**	**Ethical principle**
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.	Are back-ups made of the algorithm? Can it be restored?	
There is a much higher level of risk if there is no security by design.	Is there evidence of security by design?	2.1

Ethics[27]		
Ethical framework	**Ethical principle**	**Number**
Respect for human autonomy.	The decisions made by the algorithm are open to human checks.	1.1
Preventing damage.	The algorithm is safe and always does what it is supposed to do.	2.1
	Privacy is safeguarded and data protected.	2.2
Fairness (fair algorithms).	The algorithm takes account of diversity in the population and does not discriminate.	3.1
	The algorithm's impact on society and the environment was taken into account during its development.	3.2
Explainability and transparency.	It is possible to explain which procedures have been followed.	4.1
	It is possible to explain how the algorithm works.	4.2

Appendix 4
Endnotes

1. Statement announcing the *Understanding Algorithms* audit, Netherlands Court of Audit, February 2020.
2. For the court's ruling, see (in Dutch)
3. A *predictive* algorithm is used to analyse the question: 'What's going to happen next?' A *prescriptive* algorithm is used to analyse the question: 'What needs to be done?' (see also section 2.2).
4. The algorithm gradually discovers new interconnections (correlations) based on new data, and generates outcomes based on this. In other words, the algorithm 'learns'.
5. See audit reports published by the Netherlands Court of Audit (in 2019 and 2020). (1) *Informatiebeveiliging Verantwoordingsonderzoek 2019,* (in Dutch) (2) *Cyber security of border controls operated by Dutch border guards at Amsterdam Schiphol Airport* (20 April 2020) and (3) *Cyber security and critical water structures* (28 March 2019).
6. IT general controls (ITGC) are tools used by organisations to ensure that their IT systems are reliable and ethical. These are conventional IT tools such as those used for managing access rights (see section 4.1).
7. There are various definitions of an algorithm, all of which are more or less consistent with the wording used above. See Appendix 1 for a list of references used for this audit.
8. *Siri, Siri in my hand, who's the Fairest in the Land?*, 2018, Kaplan & Haenlein.
9. Strategic Action Plan for Artificial Intelligence, 8 October 2019, TK 2019D39726.
10. Appendix: https://www.rijksoverheid.nl/documenten/rapporten/2019/10/08/tk-bijlage-over-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid to the letter to Parliament: https://www.rijksoverheid.nl/documenten/kamerstukken/2019/10/08/tk-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid.(in Dutch)
11. In compliance with Covid-19 restrictions, only a small number of experts were allowed to attend the brainstorming session.
12. The Social Insurance Bank referred to a previous audit performed by the Netherlands Court of Audit in 2019, entitled Ouderdomsregelingen ontleed (*income Schemes for the Elderly Dissected*) (13 November 2019).
13. Appendix to letter to Parliament entitled 'Waarborgen tegen risico's van data-analyses door de overheid' (8 October 2019), TK 26643-641. (in Dutch)

14. Deep learning is a form of machine learning based on models similar to the neural networks of the human brain. Machine learning develops algorithms that allow computers to learn.
15. See our report entitled *Data-driven selection of tax returns by the Dutch Tax and Customs Administration* (11 June 2019).
16. The Control Objectives for Information and related Technology (COBIT) is an IT governance control standard designed to meet the need for assessing information-related and IT-related risks.
17. 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. Special category data is subject to additional legal protection (source: https://autoriteitpersoonsgegevens.nl/nl/over-privacy/persoonsgegevens/wat-zijn-persoonsgegevens). (in Dutch)
18. A review means that the algorithm is reassessed in order to establish whether it still complies with the relevant standards.
19. 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.
20. It is impossible to improve the operation of digital society (and the delivery of digital services) by making proper arrangements for the free movement of personal data – in other words, by making arrangements that raise and protect public confidence in society (and the government). This is the basic premise of the Dutch government's *Regie op Gegevens* ('Control of Data') programme (source: https://www.nldigitalgovernment.nl/dossiers/regie-op-gegevens-rog-control-of-data).
21. *MijnOverhei*d 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.
22. The relevant standard here is the Dutch Government Information Security Baseline, based on the international ISO/IEC 27002 standard.
23. The Information Technology Infrastructure Library (ITIL) is a benchmark for enabling IT organisations to manage operations and services.
24. Parliamentary Paper TK 35212, no. 5 dated 15 October 2020.
25. Appendix: https://www.rijksoverheid.nl/documenten/rapporten/2019/10/08/tk-bijlage-over-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid to the following letter: https://www.rijksoverheid.nl/documenten/kamerstukken/2019/10/08/tk-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid. (in Dutch)

26. The questions included in the audit framework have been formulated in part on the basis of ethical principles. The numbers refer to the ethical principles described in the table at the end of this document.
27. The questions in our audit framework are based in part on these ethical principles. Most of these principles are taken from the following European Commission reports: *Ethics guidelines for trustworthy AI* (2019) and *Whitepaper on Artificial Intelligence - A European approach to excellence and trust* (2020).

Netherlands Court of Audit
Department Communication
PO Box 20015
2500 EA The Hague
The Netherlands
Phone +31 70 342 44 00
voorlichting@rekenkamer.nl
www.courtofaudit.nl

Cover: Ontwerpwerk
Photo: Getty Images

Translater
Renée Dekker

The Hague, January 2021