nationale ombudsman

The citizen is not a dataset

Vision on the appropriate use of data and algorithms by public sector authorities

2 March 2021 | Report no. 2021/021

Research team
Emily Govers, projectleider
Danny Hanse
Gilles van Beek
Jessica Mulder
Jos van de Wiel

Text design
Xerox/OSAGE

Date: 2 March 2021
Report no.: 2021/021

Summary

Relationship between citizen and government: trust is essential

The citizen is increasingly affected by the government's use of data and algorithms. That use is not always evident and transparent. Nevertheless, it does affect the trust and confidence that the general public has in public sector authorities.

That people should be able to trust government processes is one of the fundamental tenets of our constitutional state. Trust demands more than just legal and ethical safeguards; it calls for a balanced relationship between the citizen and the government. That relationship becomes even more important in today's 'data society'.

Given the complexity of modern technology, the mechanisms underlying some government processes are becoming less transparent. They are, as it were, disappearing into a black box. At the same time, the volume of information about citizens which is collected, stored, processed and shared by government agencies is increasing. Put simply, the government knows far more about the individual. It is therefore important that, in addition to legal and ethical safeguards, our data society is subject to standards and values that are endorsed and observed by all. It falls to the government to apply these standards and values within everyday practice. This is particularly important if it is to foster and maintain the citizen's trust in its use of data and algorithms.

Clarity, accessibility and a solution-oriented approach

The government is responsible for its use of data and algorithms. (For the sake of simplicity, we use the term 'the government' to apply to all public sector authorities.) It must make all decisions relating to the use, development and application of data technology in a planned and conscientious manner. It falls to the government to ensure the quality of data-processing hardware and software, as well as the lawful and ethical use of data and algorithms. The government is responsible for all aspects of the development of the models and for their use.
This is an ongoing process which demands active input from all government agencies. Moreover, the experiences of the citizen, as well as the citizen's own wishes and requirements, should play a crucial role throughout the cycle of development, application and the use of output. Again, it is the government's responsibility to ensure that the citizen's perspective is placed to the fore, that there can be meaningful human contact, and that there remains opportunity for individualised 'bespoke' solutions where circumstances warrant. In the National Ombudsman's view, the government will fulfil its responsibility by observing three key principles: clarity, accessibility and a solution-oriented approach. These principles are discussed and further elaborated in this document.

Ombudsman's vision of appropriate use of data and algorithms by government

Offer clarity

- by identifying all use of data and algorithms, and the purpose of such use
- by determining in advance who will be involved in the processes, when and how
- by proactively offering clear information about the use of data and algorithms
- by observing legal (legislative) and ethical frameworks whereby responsibilities are duly assigned

Be accessible

- by knowing which individual the data relates to, and ensuring that he or she is able to make contact
- by involving citizens to the greatest extent possible
- by accepting and responding to questions or complaints about (the use of) data and algorithms
- by ensuring consistency and cooperation, i.e. acting as a unified government

Focus on solutions

- by determining beforehand the purposes for which data and algorithms will – and will not – be used
- by incorporating an 'emergency brake' mechanism
- by ensuring opportunity for discretion, individualisation and personal contact
- by maintaining an ongoing dialogue and by learning from any mistakes made

Lost in a digital maze

A member of the public contacted the National Ombudsman having received numerous traffic fines and road tax demands relating to vehicles that she did not own. She had contacted the appropriate authorities, which initially acknowledged the error and cancelled the charges. Nevertheless, fines and demands continued to land on her doormat. This lady requested the authorities to find a permanent solution but was told that there was nothing they could do.

The National Ombudsman contacted the relevant agencies and learned that the actual owner had given false information when registering the vehicles. It was therefore not possible to trace this person using the information in the standard registration database. Each of the three organisations concerned – the Tax and Customs Administration, the Vehicle and Driver Registration Agency (RDW) and the Central Judicial Collection Agency (CJIB) – launched its own separate investigation to find the real owner of the vehicles. They used an algorithm which, when unable to identify a particular person, would find 'possible' alternative results based on name and date of birth. This algorithm repeatedly came up with our complainant's name. Because the three organisations' databases are linked, the fines and demands continued to be sent to her address.

Although the agencies had been made aware that something was amiss, albeit at different times, they had not investigated further or modified their systems. On a few occasions, a manual check prevented a fine being issued. However, the underlying problem was not addressed. While the system was doing what it was intended to do, it was some time before anyone realised that the output was based on incorrect information.

Eventually, the issue was resolved by removing the complainant's name from the vehicles' registration details. Although a solution is welcome, it took far too long to materialise. The distress and inconvenience experienced by this lady could have been avoided. Problems like this, caused by incorrect information or errors in digital systems, are not uncommon.

The principles set out in this vision document – clarity, accessibility and a solution-oriented approach - can help to avoid this type of situation, or at least reach solutions more quickly. This case provides good example of how they will do so.

Offer clarity. The organisations concerned should have updated the data in their systems as new information became available. Even after it was evident that the data was incorrect, no attempt was made to rectify matters. It was not clear exactly who was responsible: each of the organisations worked independently of the others. By assigning responsibilities, it becomes clear who is expected to act in a situation such as this. Moreover, the system itself was not transparent. The organisations could not identify what had gone amiss or explain why the complainant had received fines in respect of vehicles that were not hers. If staff are given more complete information about the system and the data it uses, they will be better placed to spot irregularities and notify the person(s) responsible.

Be accessible. It was not easy for our complainant to contact anyone within the organisations who would be in a position to help her. She was 'passed from pillar to post' and eventually had to call on the National Ombudsman to resolve the situation on her behalf. It would have been possible for the organisations to arrive at a solution far sooner, without our intervention, had they acted together ('unified government'), removed the incorrect data and made the necessary modifications to their systems.

Focus on solutions. It would seem that too little forethought had been given to what this system could do, the situations in which it can be applied, and those in which it cannot. There appears to have been no way to interrupt the system when it produced inaccurate output, as it will inevitably do if someone has deliberately submitted false information. It should be possible for the responsible staff, perhaps on a cue from the system itself, to press an 'emergency brake', check the input and output, and implement an immediate solution. It would also be useful for the affected individual to be able to trigger a brake mechanism by notifying the authorities that something is amiss. The fact that the system continued to send out fines and demands even after the initial complaint indicates that this was not the case: the system was like a runaway train, virtually unstoppable.

This case is the subject of the National Ombudsman's report no. 2017/114, *'Verdwaald in een digitaal doolhof'*, ('Lost in a digital maze'; in Dutch).

General overview

The citizen must have trust in government action. That is one of the fundamental tenets of our constitutional state. Trust demands more than just legal and ethical safeguards; it calls for a balanced relationship between the citizen and the government. That relationship becomes even more important within our 'data society'. Given the complexity of modern technology, the mechanisms underlying some government processes are becoming less transparent. They are, as it were, disappearing into a black box.

> "Does it matter what I think? I do not approve of the government using personal information, but it does so anyway. No one has any secrets anymore: the government knows almost everything about you."
>
> **— A citizen**

At the same time, the volume of information about citizens which is collected, stored, processed and shared by government agencies is increasing. Put simply, the government knows far more about the individual. It is therefore important that, in addition to legal and ethical safeguards, our data society is subject to standards and values that are endorsed and observed by all. It falls to the government to apply these standards and values within everyday practice. This is particularly important if it is to foster and maintain the citizen's trust in its use of data and algorithms.

Take the corona test-and-trace app, for example. During its development, questions were raised regarding user privacy.[1] Third parties would be able to check whether a positive test result had been entered, which could erode public trust in the government. Besides violating individual privacy, technologies like this can create or perpetuate discrimination. The courts have already ruled that the use of the fraud detection system SyRI (System Risk Indication)[2] is in contravention of higher law due in part to an unwarranted focus on poorer neighbourhoods.[3]

The issues go beyond whether there is a justifiable reason to use data and algorithms to increase efficiency and effectiveness. There must be clarity regarding the use of data and algorithms, a careful and considered decision to do so, and due regulation of the processes which rely on data and algorithms. The government must safeguard basic (constitutional) rights such as the principle of non-discrimination, as well as procedural rights which rely heavily on the individual's

1 Verhagen, L. *Toch privacyprobleem in corona-app: patiënten zouden onder druk kunnen worden gezet.* De Volkskrant, 30 September 2020. https://www.volkskrant.nl/nieuws-achtergrond/toch-privacyprobleem-in-corona-app-patienten-zouden-onder-druk-kunnen-worden-gezet-be3d0c8b/ (in Dutch)

2 See the ruling of the District Court of The Hague, 5 February 2020, ECLI:NL:RBDHA:2020:865 (in Dutch). The court found that the enabling legislation for SyRi was not in keeping with the provisions of the European Convention on Human Rights, Art. 8 para. 2.

3 *Overheid stopt met omstreden computersysteem SyRI na uitspraak rechter,* Trouw, 5 February 2020. https://www.trouw.nl/nieuws/overheid-stopt-met-omstreden-computersysteem-syri-na-uitspraak-rechter-b62298cb/ (in Dutch)

access to information.[4] All citizens must be able to obtain clarification and explanation of matters that affect them. They must never be a 'victim' of information technology, and must be able to defend themselves against any unlawful, unfair or otherwise undesirable consequences of automated decision-making processes. Unless careful forethought is given to the use of data and algorithms, there is a risk of widespread inequality which cannot be justified or explained. For example, some ethnic groups may become subject to closer scrutiny and more frequent checks. The public must be made aware of such situations, and must be able to exert some influence where necessary and appropriate. Of perhaps even greater importance, the ongoing process of digitalisation must be accompanied by room for a personalised approach, with due regard for the wishes and requirements of the individual. It is certainly possible to achieve this when using technologies such as algorithms.

> "The tax authority fills in part of the form for you. This bothered me the first time but not now."
>
> **— A citizen**

It is important that the citizen's perspective is placed to the fore, which in the interests of good governance means that the government must involve the citizen in the entire cycle of the development and application of algorithm-based technology, and the use of its output. It must ensure that there is meaningful human contact, with room for discretion where individual circumstances warrant. The government must offer clarity, be accessible and adopt a solution-oriented approach. It is the citizen who is affected by the government's choices and actions. The output of automated processes must be in keeping with the letter and the spirit of the law. The citizen must be able to contact someone who is able and authorised to explain and, if necessary, overrule automated decisions. Government exists to serve the citizen, not vice versa.

4 Van der Sloot, B. & S. van Schendel: *De Modernisering van het Nederlands Procesrecht in het licht van Big Data: Procedurele waarborgen en een goede toegang tot het recht als randvoorwaarden voor een data-gedreven samenleving*. Tilburg University (2019), in Dutch.

Contents

1 Introduction

Government processes are being digitalised at a rapid pace. There is ever increasing use of data and algorithms. This enhances the efficiency and effectiveness of service provision, particularly where large-scale work processes can be made quicker and more accurate.[5] Increasing reliance on technology has consequences for the individual citizen, who is not necessarily aware of the process that has led to a particular action or decision.

> "If you enter various relevant factors into the computer, it saves a lot of work compared to when someone has to look everything up separately."
>
> **— A citizen**

The use of data and algorithms changes the way in which government authorities view and deal with information.[6] It creates all sorts of opportunities to simplify matters for both government staff and the citizens they serve, but computer technology is not infallible. Things can go wrong – sometimes dreadfully wrong, as the recent childcare benefits scandal amply illustrates.

> "You're placed in a group and then they subject you to extra scrutiny, as in the childcare benefits case. If you have a certain name or live in a certain postcode area, they think you're more likely to commit fraud."
>
> **— A citizen**

The National Ombudsman notes growing awareness among governmental organisations that citizens must be protected against the undesirable effects of ongoing digitalisation, and that the use of computer technology must not give rise to any form of discrimination.[7] The 'human dimension' and the assessment of cases on an individual basis remain crucial, even where the government opts to use data and algorithms to support its processes.

Constitutional rights and ethical standards must also be safeguarded. Several organisations have already introduced frameworks and guidelines, or are in the process of doing so.[8] Such frameworks are often specifically concerned with the use of data and algorithms. Although they may have no legislative basis, government organisations can commit themselves to compliance. Binding regulations governing data storage and processing also exist, as set out in the European

5 The partially completed tax return (VIA) is one example.

6 In the past, citizens were mostly assessed on the basis of their own personal information. Now they are compared to others and draw attention if they deviate too far from 'the norm'.

7 See also: *Ongevraagd advies over de effecten van de digitalisering voor de rechtsstatelijke verhoudingen.* (Raad van State, ref. W04.18.0230/I, August 2018; in Dutch).

8 In its report *Aandacht voor algoritmes* (26 January 2021), the Netherlands Court of Audit lists points for attention and improvement regarding the government's use of data and algorithms. On 10 February 2021, the Netherlands Institute for Human Rights published a handbook for (semi-) automated decision-making. It includes three basic principles which must be observed by public sector authorities in order to safeguard human rights.

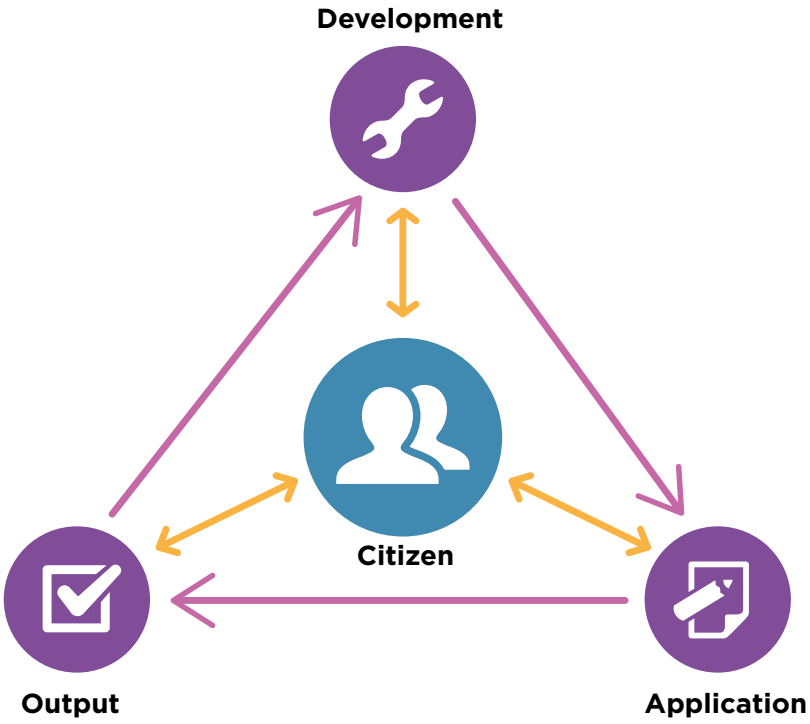
General Data Protection Regulation (GDPR) which is incorporated into Dutch law as the Algemene Verordening Gegevensbescherming (AVG).

The National Ombudsman agrees that the legislative and ethical frameworks are important, and that they should be observed in full. At the same time, we note that not all current frameworks devote specific attention to the citizen's perspective. Regardless of theme or topic, the reports and complaints received by the National Ombudsman reveal a number of recurring 'points for attention' with regard to the relationship between the citizen and the government. For this reason, the Ombudsman has formulated a number of principles which complement the existing frameworks and promote the appropriate and fair use of data and algorithms in the interests of good governance. These principles provide both a 'handhold' for government organisations and the basis of an assessment framework for the National Ombudsman.

> "I would give the government 7 out of 10. To achieve an 8, there has to be more explanation: why do authorities want to use data and algorithms and how will doing so make things better? 'Because we can' is not a good reason."
>
> **— A citizen**

2 The National Ombudsman's vision on the appropriate use of data and algorithms by government authorities

The government is responsible for ensuring the lawful, ethical and appropriate use of data and algorithms.[9] This entails an ongoing process, or cycle, in which aspects of development, application and the use of output may require modification. The process involves staff at all levels of government organisations: from administrators, managers, data scientists and developers to the front office staff who actually deal with the public. The experiences, wishes and requirements of the citizen should play a key role in all stages of the cycle.



9 There are situations in which the government will not wish to share information about the use of data and algorithms, as in matters of national security or criminal investigation, for example. Where it is not appropriate to involve the general public, the National Ombudsman nevertheless considers it important that due thought is given to the possible impact of the use of data and algorithms on individuals or groups, and that such use is lawful, ethical and appropriate at all times.

Government responsibility

The government is responsible for ensuring that the citizen's perspective is taken fully into account within the data and algorithms cycle. There must be opportunity for meaningful human contact and consideration of cases on an individual basis.
The government will fulfil this responsibility by observing the three principles of **clarity**, **accessibility** and a **solution-oriented approach**. These three principles, as formulated by the National Ombudsman, are described in brief below. A more detailed explanation follows.

Offer clarity

- by identifying all use of data and algorithms, and the purpose of such use
- by determining in advance who will be involved in the processes, when and how
- by proactively offering clear information about the use of data and algorithms
- by observing legal (legislative) and ethical frameworks whereby responsibilities are duly assigned

Be accessible

- by knowing which individual the data relates to, and ensuring that he or she is able to make contact
- by involving citizens to the greatest extent possible
- by accepting and responding to questions or complaints about (the use of) data and algorithms
- by ensuring consistency, i.e. acting as a unified government

Focus on solutions

- by determining beforehand the purposes for which data and algorithms will – and will not – be used
- by incorporating an 'emergency brake' mechanism
- by ensuring opportunity for discretion, individualisation and personal contact
- by maintaining an ongoing dialogue and by learning from any mistakes made

3 The principles in detail

Principle 1: Offer clarity

The government should provide transparency with regard to the use of data and algorithms. It bears responsibility for the entire cycle (development, application and output) and must provide information and explanation. This requires the government to identify where and when data and algorithms are being used, the purpose of such use, and the impact on groups and individuals.[10] It must maintain overall responsibility and control, even where the data and algorithms are derived from third parties (whether public or private sector).

The government must involve citizens and officials in the (proposed) use of algorithms, providing as much information as possible to those within the organisation and to the general public. It will make such information readily available to ensure that officials and citizens are aware of the situations in which data and algorithms are used, even if that use does not directly affect them.[11]

The citizen must know and understand precisely how a decision has been made, what data was used to arrive at the decision, and who is in a position to correct any inaccurate data or system errors. The staff of the various government agencies must also know how a decision, or the output of a work process, has been arrived at. Should a citizen have questions or queries, there will then be someone in a position to answer them. Last but not least, both citizens and public sector staff should be aware of the legislative and ethnical frameworks that apply to the use of data and algorithms, and who bears overall responsibility for that use. The government must act within the frameworks, and the citizen must be able to hold the government accountable.

> "I don't believe that they should be allowed to collect a huge pool of data. There must be a good reason for recording any information, established in advance."
>
> **— A citizen**

Identify the use of data and algorithms, and the purpose of such use

- During the development phase,[12] identify what data is to be used in an (algorithmic) model and why.
- Be aware of which work process(es) will rely on the output of the model and the likely impact on the citizen.
- Maintain full control of the use of data and algorithms, even if derived from third parties.

10 The greater the impact on the citizen, the more important human intervention becomes.

11 The City of Amsterdam, for example, publishes a list of the algorithms that it uses to support municipal services. See: https://algoritmeregister.amsterdam.nl/en/ai-register/

12 This is also important during the preliminary (exploratory) phase.

Determine (in advance) who will be involved, when and how

- identify the nature of their involvement.
- Identify the (types of) data in the model on which citizens, public sector officials and external experts should be consulted.[13]
- Think about how citizens, public sector officials and external experts can provide input throughout the development phase, and be able to ask pertinent questions about the use of the selected data and the model.[14]

> "I really don't know. There is so much more going on that we know nothing about."
>
> **— A citizen**

Provide clear and accessible information about the use of data and algorithms

- Provide as much information as possible about the use of data and algorithms, and do so in a proactive manner. Citizens and public sector staff should know when data and algorithms are used, even if they are not directly affected.
- Ensure that any questions about the data and model in use are answered.
- Keep the information up to date.

> "They must communicate clearly about the basis on which the machine makes decisions."
>
> **— A citizen**

Establish frameworks and responsibilities with regard to the use of data and algorithms; act accordingly

- Identify and record the responsibilities that are relevant to the use of data and algorithms at all stages of the cycle. These will include sustainable information management and the use of inclusive 'clean' data.
- Observe legislative and ethnical frameworks in all use of data and algorithms, even where the data and algorithms are derived from third parties.
- Ensure that all levels within the organisation are aware of how and where responsibilities are assigned, and that all staff act accordingly.
- Maintain an ongoing dialogue within the organisation about the use of data, algorithms and their output.

> "So what is the procedure, and how do they arrive at these decisions?"
>
> **— A citizen**

13 Especially those staff who will work with the output of the model.

14 Possible approaches include focus groups, panel discussions, customer visits and interviews with staff and experts.

Principle 2: Be accessible

The public sector infrastructure should be such as to allow all citizens full access to advice and assistance. The use of data and algorithms is intended to further this aim. With the help of staff and the general public, the government must identify the various target groups and design its processes and systems accordingly. The use of data and algorithms can create added value, e.g. in the form of 'customer profiles' linked to appropriate service provision processes. The development and application of data and algorithms should involve the citizen in a low-threshold and inviting manner.[15]

A citizen with questions or a complaint should know exactly who to contact, and should be able to contact that person in a way that he or she finds most convenient. All public sector authorities must be 'user friendly'. This applies equally to personal contact between the citizen and government staff, and at a more general level to forums such as the 'citizen panels'.

It is important that all public sector authorities act in a coordinated way – 'unified government' – and do not merely refer a citizen from one department to another. All authorities should work together to the greatest extent possible, actively attempting to solve the citizen's problem, while coordinating any necessary amendments or corrections with their chain partners (in both the public and private sectors).

> "The government uses data and algorithms to ascertain what people want, and what is required to maintain a fully functional society. The aim is to make things better and to arrive at decisions that will further the economy."
>
> **— A citizen**

Know which citizens are behind the data and be accessible to them

- Identify target groups and their requirements.
- Ask them what they need.
- Take their abilities and limitations into account.
- Ensure that no one is left behind.

Involve the citizen in the development and use of data and algorithms to the greatest extent possible

- Adopt an open attitude and actively involve the citizen.
- Invite citizens to contribute thoughts and ideas, not least with regard to the impact, direct or indirect, that the use of a particular algorithm is likely to have.
- Involve the citizen at all stages of the cycle, including the application of the model and the use of its output.

> "Machines can be used to simplify routine tasks"
>
> **— A citizen**

15 Various forms of public consultation are in use. Terms such as 'citizen panels', 'customer satisfaction surveys' and 'client councils' were used during the interviews with public sector authorities.

Be open to questions and complaints about data and algorithms; respond in a way that is appropriate to the citizen's needs

- Ensure that the organisation is accessible to anyone with a complaint or query.
- Ensure that access is easy: 'low threshold'.
- Acknowledge complaints and queries about the use of data and algorithms.
- Offer communication channels in keeping with the wishes and abilities of various citizen groups.

Act as a 'unified government'

- Always avoid referring the citizen from one department to another.
- Resolve the citizen's problem.
- Coordinate any necessary corrections or modifications with the chain partners (public or private).

> "A downside of AI and Big Data is that people no longer think for themselves. Algorithms should not make people lazy; there must be no 'dumbing down'."
>
> **— A citizen**

Principle 3: Focus on solutions

The government must ensure that any errors within a digital system[16] are promptly identified and corrected.

The government must actively monitor whether there are any errors in the system, but must also take action to preclude errors. It will involve citizens and staff, using their input to identify both known and potential errors at the earliest possible moment. This will identify situations in which the system must not be used. The potential impact on the citizen is an important consideration in this context.

The government should incorporate an 'emergency brake' mechanism so that people are not committed to systems which deliver undesirable output. It must always be possible for citizens or staff to intervene if problems occur, whereupon the citizen's situation can be assessed by other, manual means. Ideally, the system itself will identify any irregularities and alert users.

The use of data and algorithms is intended to support the work of public sector staff, not to replace them. Staff must be able to work with the automated systems while also enjoying the professional discretion to forgo their use or to deviate from the output should the situation demand. The organisational culture is an important factor in this regard. There must always be opportunity for personal contact with the citizen, and for individualised solutions.

The government must instigate an ongoing discussion about whether the algorithm and its output continue to serve their intended purpose. Both the purpose and the output must be in keeping with all current legislation. Evaluation with the help of citizens and staff will help to reveal possible errors and any necessary modifications.

> "I believe that the computer should only be used alongside a real person. A computer has no emotions, it has no empathy. When you're dealing with people, you cannot afford to ignore emotions or empathy. If you do, you will be treating them just like computers, as if they are also devoid of feelings."
>
> **— A citizen**

16 Here, 'the system' refers to any automated work process which relies on the use of data and algorithms.

Determine in advance what data and algorithms will be used for, and what they will not

- Think about what the system can actually do, and about situations in which it must not be used.
- Invite citizens and staff to contribute thoughts and ideas about the appropriate use of data and algorithms at the earliest possible stage,.
- With the help of citizens and staff, determine what is needed in cases for which the automated system is not appropriate.

Incorporate an 'emergency brake'

- Ensure that it is possible to intervene in the process and adopt alternative action where necessary.[17]
- Ensure that the system itself is able to identify irregularities and alert users accordingly.
- Ensure that some alternative solution can be offered if the system produces clearly inaccurate or undesirable output (or seems likely to do so). Establish who is responsible for that solution.

> "A computer gives you a black and white answer - there are no shades of grey. It's always 'yes' or 'no', not 'yes, but...' or 'no, perhaps...'."
>
> **— A citizen**

Provide opportunity for personal contact and an individualised approach

- Ensure that the use of data and algorithms always supports government staff rather than replacing them.
- Actively inform the public about opportunities for meaningful personal contact.
- Join the citizen in seeking a solution which is in keeping with the spirit of the relevant legislation.
- Ensure that staff can recognise undesirable output.
- Create an organisational culture in which staff enjoy professional discretion and are able to deviate from an automated decision should the situation demand.

> "If you disagree with a decision made by a machine, you feel that you've hit a brick wall. If you're dealing with a real person, you can present arguments that may influence the decision. That is the human dimension."
>
> **— A citizen**

17 Both citizens and public sector staff should be able to intervene if the system produces incorrect or inappropriate output, whereupon the situation can be assessed by other, manual, means.

Maintain an ongoing dialogue and learn from mistakes

- Ensure that there is ongoing dialogue at all levels of the organisation, involving data scientists, ICT staff, directors, managers and front office staff, all of whom should be able to 'speak each other's language'.
- Maintain and update the data and algorithms used so that the output is always in keeping with the intended purpose.
- Continually monitor output and effects, identify errors, and modify the systems accordingly.
- Regularly consult citizens and staff about their experiences and the output of the systems.

> "If I am not satisfied, I want to speak to a real person."
>
> **— A citizen**

The research process

4 Methodology

The production of the National Ombudsman's vision with regard to the appropriate use of data and algorithms by public sector authorities involved a number of research phases. Following a review of the existing literature, the researchers conducted interviews with representatives of various organisations and with experts. A public survey was conducted, the key research question of which was: "What do citizens consider important to ensure the fair and appropriate use of data and algorithms by the government?"

Desk research

The researchers studied various publications, papers and reports concerning data (including 'Big Data'), algorithms and Artificial Intelligence (AI). They also scrutinised the existing guidelines governing the use of data and algorithms, as well as relevant parliamentary papers.

In recent years, the National Ombudsman has conducted several investigations which deal with the digitalisation of government services. The publications include the Vision on Digitalisation (2017), which combines the results of several research projects and offers a number of recommendations. In 2019, an investigation examining telephone contact with public sector authorities was completed, together with a report (no. 2019/046) on the user-friendly design of online forms. An earlier report, published in 2013 (no. 2013/170), examined citizens' experiences with online government services.

The current document draws upon the insights and recommendations of these earlier publications.

Interviews

The researchers met with representatives of various public sector organisations to discuss the use of data and algorithms in their operating processes, as well as the challenges and opportunities they create. The main points arising during these interviews are reported in Section 5.

The researchers also spoke with representatives of the Ministry of Justice and Security (J&V), the Ministry of the Interior and Kingdom Relations (BZK), the Association of Netherlands Municipalities (VNG) and the Netherlands Institute for Human Rights.

In February 2021, the contents of a draft version of this document were discussed with representatives of the Employee Insurance Agency (UVW), the Social Insurance Bank (SVB), the Tax and Customs Administration, the police, and experts affiliated with the universities of Amsterdam and Utrecht.

Throughout the research process, the National Ombudsman maintained close contact with the Netherlands Court of Audit and the Council of State, both of which have a particular interest in this topic. On 26 January 2021, the Court of Audit published its own report, Aandacht voor algoritmes ('Attention for algorithms').

Public survey

The National Ombudsman commissioned the research agency TrueTalk to conduct a public survey with the key research question, "What do citizens consider important to ensure the fair and appropriate use of data and algorithms by the government?"
In-depth interviews were held with thirty members of the public, either online or by phone. The results of the survey can be found here. The quotations included throughout this report are taken from the respondents' answers (in translation).

5 Current use of data and algorithms

Introduction

Further to this research, the National Ombudsman spoke with representatives of four public sector organisations which routinely use data and algorithms in their work: the Employee Insurance Agency (UWV), the Social Insurance Bank (SVB), the police and the Tax and Customs Administration.
The objective of these interviews was to gain an impression of how the various organisations use data and algorithms. The National Ombudsman's researchers were particularly interested in the citizen's perspective: is the general public aware that the organisation uses algorithms? To what extent is the citizen affected by an organisation's data analysis practice, and are citizens involved in the development and use of data and algorithms? In this section we present the main findings with regard to the use of data and algorithms and the role of the citizen.

A framework of standards is observed

All organisations observe some form of framework governing the use of data and algorithms. This can vary from a quality framework developed by the organisation itself to one which is based on current legislation. All frameworks incorporate legislative and ethical aspects. An important component is that the organisation and the responsible staff are accountable for the use of data and algorithmic models. They must be able to justify the use of certain data within a certain model, and why this will lead to useful and useable output (the 'explainability' principle). One of the organisations is currently drawing up guidelines for communication with the public about its use of data and algorithms, and about the framework of standards it observes. Some organisations devote specific attention to the internal dialogue that must be conducted with regard to the processing of personal data, and are alert to the risk of undesirable output and its effects. To date, however, members of the public have not been involved in formulating the framework of standards.

Several organisational levels are involved

All organisations contacted by the National Ombudsman maintain an ongoing dialogue about the development, application and output of data and algorithms, in which all organisational levels are involved. This is because these organisations wish to maximise the added value of data and algorithms in their work processes and service provision. It is therefore necessary for data scientists to maintain contact with the people 'on the front line', whose experiences, wishes and requirements can then be taken into account. Moreover, these staff have direct contact with the public and are therefore aware of what citizens themselves consider important and the issues they can face. The experiences, wishes and requirements of the citizen are therefore indirectly involved in the development, application and output cycle. Some organisations directly involve the citizen, having some form of consultation mechanism such as a 'citizen panel' or 'client council'.

Data and algorithms are (thus far) primarily used to support work processes

The general impression gained from the interviews is that data and algorithms are used to support work processes. They are, for example, used to simplify routine or 'bulk' processes such as checking whether applicants meet all conditions for a deferment of payment. If so, the deferment will be granted automatically. If the application does not meet all requirements, it will be referred to a member of staff for assessment. Data and algorithms can also be used to monitor the 'digital behaviour' of visitors to an organisation's website with a view to improving service provision. Respondents also report that the use of data and algorithms can be applied to identify people who may be entitled to certain provisions that they have not yet claimed. If the various government agencies were able to share data with each other, this would further improve and expedite their service provision. At present, however, opportunities are restricted

by the current data protection (privacy) legislation. The organisations report that the use of data and algorithms helps them to identify risks and irregularities. Preliminary investigations tend to rely on data and algorithms. In addition, automated systems can identify and retrieve information that would be difficult or impossible to obtain otherwise, given the sheer volume of data involved. Staff capacity can therefore be used more efficiently and effectively. The interview respondents could not cite any cases in which the use of data and algorithms has had a direct negative effect on a citizen's situation, since such cases are always subject to manual review by a member of staff.

Results of data analyses are always checked

None of the organisations involved in this research uses the output of data analyses without applying a number of checks and balances. In all cases, the model is examined to ascertain whether the use of certain variables will influence the output and, if so, how. The types of data that can be used are restricted by current data protection legislation. One organisation reports that any 'unusual' or unexpected output is always accompanied by a warning so that staff can apply due caution. Another organisation states that it is always the official who processes a case who decides whether the output of the data analysis is likely to influence the situation of the citizen concerned. One organisation actively attempts to exclude exceptions and outliers from the system altogether, in order to avoid problems at a later stage of the process.

Not all variables are routinely used

Variables such as ethnicity, nationality and postcode are regarded as contentious and are not (or no longer) used by some organisations. All organisations apply the 'prudence principle' in their use of data. One states that it requires a minimum volume of data to support any risk assessment process. If that minimum volume is not available, the results of the risk assessment are unlikely to be representative of a larger group of people.

The citizen is not directly involved in the development and use of data and algorithms

All organisations state that the citizen's perspective is important in the development and application of data and algorithms. However, none directly involves the citizen in these phases. Various reasons for this omission are given. Members of the public lack the specific technological knowledge required; data and algorithms are used solely to support internal processes whereupon such use does not directly concern 'outsiders'. Nevertheless, some organisations do draw on information derived from consultation groups ('citizen panels'), the results of customer satisfaction surveys and the experiences of staff who have regular contact with the public. The citizen's perspective is therefore indirectly taken into account within the development, application and output cycle. The majority of organisations are prepared to consider more direct involvement, either at the individual level or through organisations which represent the interests of certain user groups.

6 Glossary and instruments

Data
Data are characteristics or information, usually in numeric form. This information is collected and collated to form 'datasets' which allow its transmission, processing or analysis. This document is concerned with data recorded electronically. The volume of data that is being recorded and stored is increasing very quickly. Public sector authorities now have a growing quantity of data relating to individual citizens. This data contains information that can be used for various purposes, including research, investigation, prevention and to make processes more efficient.

Algorithm
An algorithm is a set of instructions written in a programming language and given to a computer. The computer then performs a certain task, answers a question, makes a prognosis, supports (or actually undertakes) decision-making, or solves a problem.
There are various types of algorithm, varying from a simple 'decision tree' with a limited number of variables and pre-programmed rules, to Machine Learning (ML) and Deep Learning (DL) algorithms. ML and DL are 'self-learning' algorithms which develop autonomously without further external input and therefore fall under the heading of Artificial Intelligence (AI). These algorithms can discover complex patterns which a human would find difficult or impossible to identify. Because few people understand how they work, this type of algorithm is sometimes referred to as a 'black box'.

Automated decision-making
A decision-making process undertaken solely by algorithms, based on input data, with no human intervention.

Digitalisation
The introduction and use of computer technologies to undertake processes formerly performed manually.

Data scientist
A professional who produces computer models which undertake tasks automatically or make prognoses. A data scientist converts data into information which, in combination with knowledge and experience, can be used to derive valuable insights.

Model
A software formula within which the algorithm determines the sequence of process steps.

Variable
A characteristic whose value can differ, e.g. age, income, social class, etc.

Big data
Big Data is a term used to refer to a massive volume of both structured and unstructured data that is so large it is difficult to process using traditional database and software techniques. Such data is also extremely diverse, and can include text, numbers, images, videos and audio recordings. Big Data is the result of the trend of collecting and storing an ever increasing quantity of data, accompanied by ongoing improvement in the processing power of computers and the development of new software. The analysis of Big Data relies on advanced algorithms and/or Artificial Intelligence.

What instruments can support the appropriate use of data and algorithms?

The use of both data and algorithms raises some ethical questions. Instruments that can help ensure the appropriate and fair use of data and algorithms include DEDA, BIAS and the FairTrade method. In all cases, it is important to take the citizen's perspective into consideration, insofar as the instruments themselves do not already do so.

DEDA

DEDA (Ethical Data Assistant) is an instrument which helps to identify ethical issues, develop an awareness of value conflicts within a data project, or document the ethical decision-making process. The instrument consists of a manual, an app and a poster intended to promote open discussion. It is important to consider all listed ethical perspectives (moral relativism, utilitarianism, etc.). The use of only one perspective will lead to different results for different citizens because aspects such as culture or membership of a minority group are overlooked. DEDA was developed using an iterative process and in close consultation with public sector officials, academic researchers and members of the public. One limitation is that the latter did not form a representative sample of the Dutch population. The developers nevertheless attempted to maintain a balance between a workable development process and the input of as many stakeholders as possible.

BIAS

BIAS (a 'backronym' formed from the Dutch for 'decision-making instrument for algorithmic systems') is an instrument that helps to streamline ethical consideration and accountability processes. It is still in development. The focus is on public values, i.e. which values could be damaged and which achieved or enhanced, the choices and balances that must be in place to ensure a favourable outcome, and how the values are to be safeguarded. The developers contend that the citizen's perspective is closely allied with these public values. As yet, members of the general public have not been actively involved in the development process.

FairTrade method

The FairTrade method can be used to promote fairness within ML models, thus preventing discrimination. The method reveals whether a model is intrinsically fair and unbiased, the requirements that datasets must meet, the requirements that the operating team must meet, and the requirements that communication and decision-taking must meet. Various steps are undertaken in order to establish the relationships between variables (using hypotheses and a combination of hypotheses and data in order to assess the form and strength of the relationships). Variables which are shown to have undue influence on other variables, thus leading to biased output, are removed from the model. Next, a model is trained using only 'pure' unbiased information. When formulating hypotheses about the relationships between variables, specific attention is devoted to (the absence of) bias, because it is at this stage that the risk to the citizen is greatest.

DEDA: https://dataschool.nl/deda/
BIAS: https://dataschool.nl/samenwerken/bias/
FairTrade method: CBS

Nationale ombudsman
Postbus 93122
2509 AC Den Haag

Telephone 070 356 35 63
nationaleombudsman.nl

Report no. 2021/021 is a publication of the
Nationale ombudsman, March 2021