HRIA on government Facebook Pages

Human Rights Impact Assessment on the processing of personal data on government Facebook Pages

Colofon

Date	18 November 2022
Status	Public version
DPIA by	**Ministry of the Interior and Kingdom Relations** Turfmarkt 147 2511 DP The Hague PO Box 20011 2500 EA The Hague **www.rijksoverheid.nl/bzk**
Contact	Press department Ministry T: +31 70 426 88 88
Project Name	**HRIA report** on the processing of personal data on government Facebook Pages
Authors	Winfried Tilanus https://www.privacycompany.eu/

Contents

Summary

This report, commissioned by the Dutch Ministry of the Interior and Kingdom Relations, assesses the human right risks of the use of Facebook Pages by the Dutch government. This Human Rights Impact Assessment (HRIA) focusses on the risks for data subject's rights to non-discrimination, freedom of thought, conscience and religion, and freedom of expression and information. This HRIA must be read together with the (extensive) Data Protection Impact Assessment (DPIA) performed on the same data processing.

This assessment concludes the use of Facebook Pages by Dutch government organisations has a potentially high impact on human rights. Due to the lack of transparency (including meaningful access to Facebook's actual big data processing) the real impact could not be assessed. Dutch government organisations do not have the means to assess, control, correct/minimise or explain the algorithms used by Facebook in relation to visits to the content they publish on Facebook Pages. Absent transparency and means to minimise the potential negative impacts on the persons visiting the government Pages, government organisations must assume the use of Facebook Pages has a high risk for human rights.

When a person visits a Dutch government Facebook Page, that visit may change the personalisation of the timeline and the advertisements shown to that person. All behaviour on Pages is stored by Facebook, such as opening a post, time spent reading a post, liking it and/or forwarding it to a friend. Because there is too much information on Facebook for users to scroll through all updates from their friends, the friends of their friends, and content from followed Pages, Facebook deploys algorithms to select the content it presents to its users. This HRIA describes six potential sources of bias, and concludes that bias by one sided suggestions is present and that biases by feedback loops and bias by optimalisation criteria are very likely to be present. This conclusion is based on the observations of the suggestions Facebook made during the (limited) research for this DPIA and HRIA: the algorithms selected anti-governmental opinions that attract the most likes. The initial criterium of suggesting the posts with the most likes already seems to introduce a bias. On top of it, it steers into one direction. Over time this may reinforce itself because the highest suggested posts are the most likely to receive new clicks and likes.

The algorithms may steer users towards less diversity or representation of minorities, push users towards certain actions, opinions or lines of thought, while discriminating/segregating other groups.

Facebook did not provide exhaustive information on the data it uses for its algorithmic decisions, and did not provide insights in the logic either in reply to individual data subject access requests. This lack of transparency makes it impossible to assess what impact the personalisation of Facebook has on human rights when the government uses Facebook Pages.

Foreign, non-EU governments can impact human rights of the persons visiting a Dutch government Facebook Page. Government authorities may compel disclosure, use OSINT or use ADINT to monitor individual visits to government Facebook Pages. They may also compel Facebook (or hack/bribe Facebook employees) to disclose data it has inferred from those visits. This may result in intimidation or (cyber)attacks against individual visitors of a government Page or may result in human right violations when such people travel to third countries. As described in the DPIA, it follows from the Schrems-II jurisprudence of the European Court of Justice there is a real risk that government authorities in the USA can access personal data collected by Facebook. That is why the Irish data protection commissioner has issued a draft ban on future data transfers from EU Facebook customers to the USA.

This HRIA concludes that government organisations should refrain from using Facebook Pages as a communication medium, in view of the high risks for human rights of the Page visitors. This may change once new rules from the EU Digital Services Act enter into effect for Facebook as a Very Large Online Platform, but Facebook’s compliance with these rules will have to be assessed.

1. Introduction

This report, commissioned by Ministry of the Interior and Kingdom Relations, is a Human Rights Impact Assessment (HRIA), on the processing of personal data on government Facebook Pages. This HRIA focusses on the risks for data subject's rights to non-discrimination, freedom of thought, conscience and religion, and freedom of expression and information. It supplements the Data Protection Impact Assessment (DPIA) on the same data processing and draws on the same research as performed for the DPIA. For the sake of brevity, this report refers to the DPIA where possible.

In October 2021, Facebook changed its corporate name to Meta. In this report 'Facebook' will be used for the social media platform, to prevent confusion with other apps offered by Meta such as Instagram and WhatsApp.

1.1 Facebook Pages

Formerly Facebook Pages for organisations were known as *Fan Pages*. Government organisations can and want to use Facebook Pages to reach a broad audience. Facebook enables government organisations to directly communicate with people in a way they are used to, through the platform where they already spend a lot of time.

A Facebook Page from a government organisation can be viewed by both Facebook users and non-Facebook users. Facebook users who like or follow a Page will get updates from that organisation in their *News Feed*. Even if they do not follow the government Page, they may see a recommendation if their friends follow the Page, or like a post on such a Page.

1.2 Human Rights Impact Assessment

A Human Rights Impact Assessment assesses the human rights impacted by the deployment of a certain system in a given context. One of the methodologies to do so, is a model for the assessment of the human rights impact of algorithms, 'Impact Assessment Mensenrechten en Algoritmes' (IAMA).[1] This model is part of a toolbox for ethical responsible innovation, 'Toolbox Ethisch Verantwoorde Innovatie'[2], published by the central Dutch Government. This IAMA model focusses solely on human rights impacts caused by algorithms and not on human rights impacts caused by other aspects of a system. This HRIA follows the model of the IAMA, but expands it to other mechanisms that may cause a human rights impact, such as third country government access to personal data processed by Facebook as a result of visits to a government Page.

The IAMA model has a structured set of questions, divided in 4 groups:

The purpose of the processing
The characteristics of the algorithm and the data used by it
The implementation of the algorithm and the use of its output
The human rights impacted by the algorithm

[1] Impact Assessment Mensenrechten en Algoritmes (IAMA);
https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes

[2] Toolbox Ethisch Verantwoorde Innovatie;
https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/nieuwe-technologieen-data-en-ethiek/publieke-waarden/toolbox-voor-ethisch-verantwoorde-innovatie/

This HRIA follows these questions. Because the IAMA model is in Dutch, the questions and some other parts of the model are translated. Further explanations and the legal background can be found in the IAMA model.

1.3 Human Rights

Human rights are safeguarded in several treaties and laws.[3] This HRIA mainly focusses on the European Convention of Human Rights[4], but also refers to the Charter of Fundamental Rights of the European Union[5] ('Charter').

The European Convention of Human Rights is applicable law within the 46 Member States of the Council of Europe. The Netherlands is one of these Member States. Anyone who feels these rights are violated by a Member State, can directly take it to the European Court of Human Rights. Within the EU, fundamental rights are guaranteed in Charter of Fundamental Rights of the European Union. This Charter covers the European institutions and the implementation of European regulations in national law. The European Court of Justice interprets the Charter. These two treaties are important for this HRIA because they precede above national law.

The IAMA Model summarizes the human rights from the different treaties to the following groups of rights:

Rights appertaining to the person
1 Personal identity / personality rights / personal autonomy
2 Social identity / relational privacy rights / relational autonomy
3 Physical and mental integrity
4 Data protection / informational privacy rights
5 Communication rights
6 Spatial privacy rights
7 Property bound privacy rights
8 Reputation Rights (The right to protection of reputation)
9 Healthy living environment (Environmental protection)
10 Social and economic rights

Freedoms
11 Freedom of expression
12 Freedom to receive information
13 Freedom of religion
14 Freedom of assembly and protest
15 Freedom of association
16 Political rights/freedoms

Equality rights
17 Equality before the law
18 Prohibition of direct discrimination on certain grounds
19 Prohibition of indirect discrimination on certain grounds
20 Prohibition of discriminatorily motivated action
21 Right to material distinction / customization
22 Right to reasonable accommodation / affirmative action
23 Prohibition of profiling
24 Prohibition of segregation

Procedural rights

[3] https://www.rijksoverheid.nl/onderwerpen/mensenrechten/mensenrechten-nederland
[4] https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=005
[5] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012P/TXT&from=EN

25 Right to proper administration (pre-trial stage)
26 Right to an effective remedy and access to justice (trial phase)
27 Right to an independent and impartial judge
28 Right to a decision within a reasonable time
29 Right to a fair trial
30 Right to a fair trial: criminal justice safeguards
31 Criminal legality requirement (no punishment without law)

The HRIA will follow this list of human rights from the IAMA model, when analysing the human right impact.

Both the European Convention of Human Rights (art. 14)[6] and the Dutch constitution (art. 1)[7] prohibit discrimination but neither of them limits the possible grounds for discrimination. At the same time making distinction on any of these grounds is only discrimination when it has "no objective and reasonable justification"[8]. So it depends on the context if making a distinction on a certain ground is discrimination. When analysing possible discriminatory effects of a system, it is needed to first determine what grounds for making distinction is discriminatory within the context of the deployment of that system. This HRIA uses the term 'protected properties' for all grounds for discrimination within the context of the system that is analysed.

1.4 Scope of this HRIA

This HRIA report assesses the impact on human rights when a user visits a Facebook Page created by a Dutch government organisation. The scope includes both registered users of a Facebook Page, and visitors to a government Page that do not have a Facebook account (non-users).

The scope includes the collection of off platform data about non-users (with cookies) as these people may seek government information that is only available on Facebook, or inadvertently visit a public Facebook page as a result of a search query without having accepted Facebook's terms and conditions. This type of data processing is in scope because the data processing (with the cookie) originates from a visit to a government Page. This is relevant processing of personal data of persons visiting the Page.

The scope includes the processing necessary for Facebook to show recommendations to visitors of the government Facebook Page, including recommendations created by Facebook's algorithms to rank content based on inferred preferences. This is in scope because the data processing occurs on a government Page and is the result of algorithmic processing of personal data of persons visiting the Page.

As described in the DPIA, for the purpose of this DPIA and HRIA a test Page was created, the (fictive) Ministry of Privacy.

This HRIA depends on the research and analysis of the interactions with the test Page described in the DPIA on government Facebook Pages.

2. Analytical framework

To fully understand the potential impact on human rights of the use of Facebook Pages by government organisations, the IAMA model needs to be supplemented with some theoretical background

[6] https://www.echr.coe.int/Documents/Convention_ENG.pdf
[7] https://wetten.overheid.nl/jci1.3:c:BWBR0001840&hoofdstuk=1&artikel=1&z=2018-12-21&g=2018-12-21
[8] See for example Case of J.D. and A v. the United Kingdom, ECLI:CE:ECHR:2019:1024JUD003294917

First, human rights are linked to personal identity and autonomy. On an intuitive level, it is evident that there is a relation between identity and autonomy on one hand and personalisation on the other. But to assess the potential impact of personalisation on personal identity and autonomy, a theory about the relation between these concepts is needed. The first part of this analytical framework offers such a theory.

Secondly, biases in algorithms can have a big impact on human rights. But it is hard to assess whether an algorithm is biased or not. So before assessing the impact on human rights, it is necessary to know where these biases in algorithms originate from, how they can be detected and how the impact can be mitigated. The second part of this analytical framework discusses different types of bias in algorithms.

Thirdly, other potential human rights violations originate from the global accessibility of information published on Facebook. Authorities and companies in third countries can collect information about visits to a Facebook Page of a government organisation and use this information against inhabitants of Europe. The third part of this analytical framework discusses different ways of analysing such extraterritorial human rights violations.

2.1 Personal autonomy, identity and personalisation

To understand the relation between personal autonomy, identity and personalisation, this HRIA uses the analysis Stuart Hall made of the concept ‘cultural identity’.[9] Stuart Hall breaks ‘identity’ down into two conflicting processes, a Foucaultian process of imposing identities on people and a Freudian process of self-identification. The Foucaultian process is a process of executing power, the Freudian process is the source of human agency.

Human autonomy and human identity development are directly linked to this Freudian process of self-identification. Self-identification always is a process of identifying yourself with somebody or something you recognize yourself in. This can only happen when there are examples available to identify yourself with, in the personal life or publicly visible in the media. So key to human autonomy and to human identity development is the visibility of diverse groups and minorities in both the personal life and the media.

A social network potentially offers a platform where minorities can be visible and represented. At the same time the personalisation of a social network can vastly limit the diversity of people becoming visible. This can even result in a Foucaultian process of imposing an identity upon somebody. For example when the personalisation algorithm of the social network has a strong preference to present women in their late twenties with newsfeeds and images about families with young children, then these women get the identity of ‘married mother’ imposed upon them. Other identities, like unmarried career woman, disappear from the representation. This hinders the ability for women in their late twenties to identify with other identities then ‘married mother’ and so limits their ability form their own identity and limits their ability to make autonomous choices.

2.2 Bias in algorithms

Visits to a government Facebook Page can have an impact on some human rights when the algorithms Facebook uses for personalisation have a bias. Bias is an inclination or prejudice for or against a person or group, especially in a way that is considered to be unfair.[10] To understand the potential human rights impact of the deployment of algorithms, it is important to recognize

[9] Stuart Hall; Who Needs ‘Identity’? in: Questions of Cultural Identity; Edited by Stuart Hall and Paul du Gay (London, 1996).

[10] David Marshall; “Recognizing your unconscious bias,” Business Matters, www.bmmagazine.co.uk/in-business/recognising-unconscious-bias/ (October 22, 2013)

potential sources for bias. For this HRIA the analysis of the data science processes, algorithms and bias by Tobias Baer is used as analytical framework for bias in algorithms.[11]

This HRIA does not make a clear distinction between algorithms that are composed by data scientists and algorithms that make use of machine learning. Machine learning algorithms can incorporate more diverse and more complex data then data scientists are able to process, but both methodologies are prone to the same sources for bias and the same systematic errors. Machine learning can be more sensitive to some of the potential sources for bias, because of the large amounts of processed data and the lower level of human oversight.

2.2.1 Sources for bias in algorithms

Both when an algorithm is fully developed by data scientists or when data scientists use machine learning, the developers need data to determine what choices the algorithm should make. When these data contain a bias, then this bias is replicated in the algorithm, resulting in biased outcomes of the algorithm. Biases have multiple origins. Below, six origins are explained in a bit more detail.

First of all, in our society there are many prejudices, based on for example sex or age. These prejudices will be included in the data used for developing the algorithm and because of that inclusion, the algorithm will be biased too. Countering this bias means countering the prejudices in our society. An example of such a prejudice is that women become mother when they reach a certain age. This prejudice may create a bias in the algorithm, resulting in a preference for motherhood related items in suggested content to Facebook users. To counter this bias, the algorithm could deliberately suggest items related to woman without children. Countering this type of bias is a political decision.

Secondly, the data used for developing the algorithm lag a bit behind the changes in society, resulting in algorithms being a bit biased towards maintaining the status quo, as opposed to following new trends. This bias can be countered by constantly feeding an algorithm with new data and designing it to pick up new trends.

Thirdly, the algorithm can become biased when there is a feedback loop between the outcomes of the algorithm and the data used create the algorithm. This happens for example when a list of suggestions is ordered based on the amount of clicks on each item. When the algorithm is updated based on the current amount of clicks on each item, then there is a feedback loop, because the list was already ordered and people tend to mainly click on the topmost items. This results in a bias that amplifies a small trend more and more for each iteration of the algorithm. When the ordering of suggestions is done on an individual level, each user may be pushed more and more into their own bubble. This can only be countered by collecting the dataset for de development of an algorithm totally independent from the deployment of the algorithm.

When the datasets used to create the algorithm are free from bias, usually an algorithm has the same outcome for similar input. However, when an algorithm is sensitive for somebodies situation or some properties of that person, it will direct most of its suggestions in one direction. This is a fourth source of biases. Making suggestions in one direction can cause the creation of a ‘bubble’: an environment that has a tendency towards a certain topic or opinion while making other topics and opinions invisible. During the research for the DPIA and this HRIA this type of bias was visible because test user B followed all political parties that had a Facebook page, but the News Feed of test user B showed an increasing amount of anti-government messages (DPIA, paragraph 1.1.3 personalisation). This type of bias can be countered by deliberately adding suggestions about different topics and showing different opinions. In practice this is hard to do, because the algorithm still has to choose which of the other topics or opinions it will show.

[11] Tobias Baer; Understand, Manage, and Prevent Algorithmic Bias: A Guide for Business Users and Data Scientists (Kaufbeuren, Germany, 2019)

Data scientists must make choices about the criteria used for optimisation when designing an algorithm. When for example the amount of clicks is used to determine the ranking of a link, then links with catchy, click-bait like, titles will be preferred above more descriptive links. This will cause the algorithm to start preferring click-baits and rage-baits. When the quality of a suggestion is defined as 'it is outside the usual of the user, but the user still has chosen it', then the algorithm will prefer change. Because most people have a moderate opinion in the middle, such an algorithm would start pushing users towards more extreme opinions on either side. This type of bias is fundamental to algorithm design and countering this bias can only be done by making it part of the design process of the algorithm in its wider context.

Finally, during the development of an algorithm data scientists have to make smaller choices that can cause bias. As part of the normal process of preparing data for use, data are cleaned-up. That means 'outliers' (points that have extreme deviating values) are deleted. But when does a point have extreme deviating values? The data scientists have to decide what points to keep and which ones to throw away. With such decisions data scientists can introduce new bias.

It is also common practice to finetune the algorithm based on a review of the outcomes. In this process the data scientist has to decide what optimisations are correct or not, again possibly introducing bias. Countering this type of bias is hard, because there can always be a bias in the choices of data scientists. Careful choices by data scientists who have good insight in both the algorithm and the context of its deployment might help here, as may public accountability for these choices.

2.2.2 Direct and indirect discrimination

Direct discrimination happens when the property that is used to treat two people differently, e.g. gender or a physical trait common among people of a certain ancestry[12], is registered directly and used in the algorithm to make the distinction. Indirect discrimination happens when other information that does not contain the property itself but is strongly related to such a property is used to make the distinction. An example of such indirect discrimination, or discrimination *by proxy*, is when groups of people are treated differently based on the zip code of their residence. Because some residential areas are more or less segregated by income or physical traits, use of the zip code results in indirect discrimination based on these characteristics. Data scientists can try to use their knowledge to discard all variables that can cause indirect discrimination. When using machine learning, this becomes more complicated:

> *"... if the bias is already present in the data used to train the model (e.g., because it mirrors societal biases), the machine learning algorithm will go out of its way to capture indicators for the bias. If you remove direct indicators (e.g., the ZIP code), it will find indirect ones (e.g., the distance to Joe's Potato). If you remove the indirect ones, it will find even more indirect ones (e.g., the number of businesses with a name starting with "J" in the vicinity of the applicant). This is why it is not always possible to remove the bias from the model ..." (Bear, p.156)*

Indirect discrimination can be detected with statistical analysis. To do so, the protected attribute (e.g. a physical trait common among people of a certain ancestry) and the dataset or the outcome of the algorithm need to be correlated. This protected attribute is often not collected. It might even be prohibited to collect it.[13] Without access to the protected attribute it is virtually impossible to detect indirect discrimination.

[12] The term 'race' or 'racial origin' is not used, to prevent the suggestion of acceptance of theories which attempt to determine the existence of separate human races. See Recital 51 of the GDPR.

[13] Based on the general prohibition on the processing of special categories of personal data in the GDPR (General Data Protection Regulation), article 9.

2.2.3 Judging bias

Judging to what extent an algorithm is biased can either be done by tight control and rigorous scrutiny while developing the algorithm or by observing the behaviour of the algorithm. In both cases it must be clear beforehand which types of bias are undesirable and which types of bias are wanted. This choice is a political decision, though human rights already provide a good starting point. To assess the algorithm it should be assessed in isolation from the systems and user interfaces it is connected to. The algorithm should be run repeatedly in a test environment with small changes to the input data. Access to the underlying data, inferences and decisions is also needed when analysing if the dataset contains properties that are indirectly discriminatory. Without such a laboratory setup, it is impossible to fully understand the behaviour of an algorithm.

2.3 Extraterritorial Human Rights violations

The most comprehensive way to enforce human rights within the EU, is by appealing to the European Convention of Human Rights and the Charter of Fundamental Rights of the European Union.

Because Facebook operates on a global base, data processed by Facebook can be accessible in countries that are not member of the Council of Europe. This global accessibility can impact the human rights of visitors of Dutch government Pages in several ways:

- When a visitor of a Dutch government Page travels to a country that is not member of the Council of Europe, that visitor might be prosecuted according to local criminal law. For example a man who has publicly liked government content messages about homosexuality can be prosecuted in a country where practising homosexuality is prohibited.

- When a hostile state actor from outside the Council of Europe directly stalks, intimidates or (cyber)attacks a visitor of a Dutch government Page, for example an attack against a dissident who is seeking refuge in the Netherlands and is interacting with a Page from the IND (Immigratie- en Naturalisatie Dienst).

- When third countries impose rules or norms on Dutch visitors of a government Facebook Page in a way that violates human rights, for example by not suggesting posts that would be protected by freedom of speech rights according to the European Convention on Human Rights. For example by removing government content on sexual health from a timeline because that is regarded ‘indecent’ in other cultures.

Government authorities in third countries can use Facebook to gather the intelligence needed to prosecute or attack visitors of Dutch Facebook Pages. This can be done in several ways:

- By compelling disclosure of personal data stored outside Council of Europe member states or by laws with an extraterritorial scope. These data protection risks are described in the DPIA.

- By Open Source intelligence, for example when the border police is checking what government Pages a traveller follows when crossing a countries border.

- By ‘Advertisement Intelligence’ (ADINT)[14]. ADINT enables intelligence services and other state actors to deduce personal aspects of people through the advertisement interfaces of Facebook. These interfaces may leak sensitive information. The information Facebook stores and infers can include information about viewing or following a government Facebook page.

[14] https://adint.cs.washington.edu/, ADINT is for example used by the Dutch intelligence services AIVD and MIVD, see: https://www.ctivd.nl/documenten/rapporten/2022/02/08/rapport-74

Through these mechanisms some of the human rights of visitors of Dutch government Facebook Pages can still be impacted by non-member states.

Assessment according to the IAMA model

3. Why – the envisioned effects, purposes and objectives

Part 1 of the IAMA model deals with the "Why?" of the intention to develop, procure, adapt and/or deploy an algorithm (hereafter, for brevity: algorithm deployment). What are the reasons, the underlying rationale and the intended effects of the algorithm deployment? What are the underlying values that determine the approach to using the algorithm? These overarching questions should be addressed first in a decision-making process about the deployment of algorithms, before, for example, getting to questions about preconditions or possible impact on fundamental rights.

3.1 Reason and problem definition

Explain your proposal for the use/deployment of an algorithm. What has been the reason for this? What problem should the algorithm solve?

In order to properly function, governments need to actively communicate with the inhabitants of their territory. The reach can be vastly improved when the communication is sent via a platform that allows for push messages and via a platform with a big audience. Facebook is such a platform. The amount of information on Facebook is too big for any single user to easily process. Facebook's algorithms help users to prioritise content on their timeline. The algorithms also help by providing suggestions for other content.

3.2 Purposes

What purpose is to be achieved by the deployment of the algorithm? What is the main purpose and what are sub purposes?

Facebook's purposes of the data processing are described in detail in the DPIA (Section 4, Purposes of the processing). The purposes of government organisations to open a Facebook Page have been described in the reply to the first question.

The algorithms are used for personalisation. The DPIA describes several types of personalisation in relation to the government Facebook Page:

1. Postings shown in the *News Feed* from the specific people and organisations followed by the two test users including the government test Page.
2. Postings shown in the *News Feed* based on Facebook's algorithmic recommendations in relation to the daily visits to the government test Page.
3. Advertisements shown as 'sponsored content'
4. Advertisements shown as 'sponsored posts'
5. Advertisements shown as 'related pages' when visiting the government test Page

These personalisations are determined by algorithms.

3.3 Public Values

What are the public values that inform the deployment of the algorithm? If multiple public values inform the deployment of the algorithm, can a ranking be applied to them?

Facebook and its algorithms are deployed to make the vast amounts of information, including the government communications, accessible for existing and potential Facebook users. This value directly drives the human right of freedom to receive information.

As a side effect, the use of Facebook may support:

1. Personal identity / personality rights / personal autonomy
2. Freedom of expression
3. Freedom of assembly and protest

(see part 4 of this HRIA for a detailed analysis)

What are the public values that might be compromised by the deployment of the algorithm?

Of the 31 human rights listed in the IAMA model, this HRIA finds that the use of government pages potentially has an impact on the following 21 human rights (see Section 6.2 for a detailed analysis):

1. Personal identity / personality rights / personal autonomy
2. Physical and mental integrity
3. Data protection / informational privacy rights
4. Communication rights
5. Spatial privacy rights
6. Property bound privacy rights
7. Freedom of expression
8. Freedom to receive information
9. Freedom of demonstration
10. Prohibition of indirect discrimination on certain grounds
11. Prohibition of discriminatorily motivated action
12. Right to reasonable accommodation / affirmative action
13. Prohibition of profiling
14. Prohibition of segregation
15. Right to proper administration (pre-trial stage)
16. Right to an effective remedy and access to justice (trial phase)
17. Right to an independent and impartial judge
18. Right to a decision within a reasonable time
19. Right to a fair trial
20. Right to a fair trial: criminal justice safeguards
21. Criminal legality requirement (no punishment without law)

3.4 Legal Basis

What is the legal basis of the deployment of this algorithm and of the intended decisions that will be made based on this algorithm?

See Section 11 of the DPIA. Facebook explains to its users in its general Privacy Policy (last updated 26 July 2022) that for personalisation it relies on the legal ground of the necessity to perform a contract.[15] There are two exceptions: Facebook relies on consent for cookies to personalise ads on and off Facebook, and Facebook relies on explicit consent when it uses special categories of data actively provided by users, Facebook.[16] This legal ground of explicit consent does not apply to any observed or inferred special categories of data. The DPIA concludes that Facebook does not have a legal ground for the processing of inferred sensitive data.

[15] Meta Privacy Center, Privacy Policy, What is our legal basis for processing your data, URL: https://www.facebook.com/privacy/policy/?section_id=7-WhatIsOurLegal. Facebook describes that this legal ground includes:

- Personalise features and content (such as your News Feed and Stories);
- Personalise the ads people see, and
- Make suggestions for you (such as people you may know, groups or events that you may be interested in or topics that you may want to follow) on and off our Products

[16] Meta Privacy Policy, last updated 26 July 2022, What is our legal basis, URL: https://www.facebook.com/privacy/policy/?subpage=7.subpage.1-WhatIsOurLegal

Similarly, the DPIA concludes Facebook does not have a legal ground for the setting and reading of datr tracking cookies when users or non-users visit a government Page and they do not provide consent to marketing cookies.[17]

3.5 Stakeholders and responsibilities

What parties and individuals are involved in the development/use/maintenance of the algorithm?

Facebook develops, deploys and maintains the algorithms. The DPIA describes that admins and users cannot actively influence or minimise the data processing. Users can only 'feed' the algorithm by following/unfollowing Pages, liking/unliking content and clicking on links, but such actions do no directly influence the ranking in the algorithm. Administrators of government Pages can prevent Facebook from showing other recommended Pages, but do not have an option to exclude visits to a government Page from the personalisation algorithm.

How are the responsibilities allocated with respect to the development and deployment of the algorithm, even after the algorithm is completed?

The algorithms are developed and deployed under the responsibility of Facebook. Facebook legally offers a joint controller agreement for the creation of statistics about the visits to a government Page. Facebook hence unilaterally divides responsibility for this specific output of the algorithmic data processing. The DPIA describes that this assignment is incomplete, as Facebook continues to independently process the collected personal data about visits to government Pages for its own purposes, to 'feed' its ranking algorithm and personalise contents.

Who is ultimately responsible for the algorithm?

See Section 5 of the DPIA, Processor or (joint) controller.

4. The Algorithm

Once it is determined why an algorithm will be deployed and how the organizational safeguarding of public values and interests will take shape, it is important to think about the design of the algorithm to be deployed. This is what this part of the IAMA model is about, which deals with the "What?" of the project. This part is divided into two parts: part A concerns the input of the algorithm: the data (or digitally captured data) that will be used and the boundary conditions around it. Part B concerns the algorithm itself, that is, the throughput of the project. Like the previous part, This part includes a number of questions and points that should be considered in any decision-making process on the deployment of an algorithm.

4.1 Part A Data – Input

4.1.1 Estimation: type of algorithm

What type of algorithm is being used, or what type of algorithm is going to be developed?

Facebooks deployment consists of a cluster of algorithms/tasks that interact with its data storages (Hive and other databases) and with each other. Facebook deploys both self-learning algorithms and non-self-learning algorithms. Facebook regards the exact inner workings of its systems a trade secret. For more details see Section 8 of the DPIA.

[17] In the description of its 'legitimate interest' purposes, Facebook includes a category of users named '*If you are using a device we cannot associate with a registered user of the Meta Products*'. Facebook does describe that it relies on this legal ground for the processing of cookies, but does not include the purpose of behavioural advertising.

4.1.2 *Data sources and quality*

What type of data will be used as input to the algorithm and from what sources does the data come from?

Facebook regards the exact inner workings of its systems a trade secret and does not publish what input is used for each algorithm.

Is the quality and reliability of the data sufficient for the intended data application? Explain.

Facebook regards the exact inner workings of its systems a trade secret and does not publish information about the quality of the data used.

4.1.3 *Bias and assumptions in the data*

What assumptions and bias are embedded in the data and how is their influence on the output of the algorithm is corrected or otherwise overcome or mitigated?

Facebook regards the exact inner workings of its systems a trade secret and does not publish information on biases in the data, and does not give access to its customers or independent scientific researchers to investigate bias and assumptions in the data.[18] Facebook does make limited data available to researchers in an archive of (political) advertising (as do Twitter and Google). However, according to academics from the university of Amsterdam, these archives *have faced difficulty in defining and identifying, at scale, what constitutes a "political advertisement". Secondly, verifying: ad archives have proven vulnerable to inauthentic behaviour, particularly from ad buyers seeking to hide their true identity or the origin of their funding. Thirdly, targeting data: ad archives do not document in meaningful detail how ads are targeted or distributed*."[19] The dataset with political ads Facebook provides, is incomplete because advertisers have to report themselves when an advertisement is a political advertisement. This results in an underreporting of political advertisement.

In the next few months, when Very Large Platforms have been formerly identified that must comply with specific transparency rules in the EU Digital Services Act (DSA), Facebook will have to publish an ad inventory.[20] This inventory will have to specify the parameters used to select the groups of recipients and parameters used to exclude one or more of such particular groups.[21]

---

[18] Medium, Researchers, NYU, Knight Institute Condemn Facebook's Effort to Squelch Independent Research about Misinformation, 4 August 2021, URL: https://medium.com/cybersecurity-for-democracy/researchers-nyu-knight-institute-condemn-facebooks-effort-to-squelch-independent-research-about-59cec0793939

[19] Joren Vrancken, Theme Analysis of Political Facebook Ads in the 2021 Dutch General Election, January 2022, URL: https://www.researchgate.net/publication/357790703_Theme_Analysis_of_Political_Facebook_Ads_in_the_2021_Dutch_General_Election. Zie ook: Ausloos, J., Helberger, N., Leerssen, P., Vreese, C.H. de, Zarouali, B, Platform ad archives: promises and pitfalls, in: Internet Policy Review, vol. 8, no. 4 2019, URL: https://policyreview.info/articles/analysis/platform-ad-archives-promises-and-pitfalls

[20] The DSA was adopted by the European Parliament on 5 July 2022, consolidated text at URL: https://www.europarl.europa.eu/doceo/document/TA-9-2022-0269_EN.html#title2. Article 30 of the DSA contains an obligation for Very Large Platforms such as Facebook to publish a registry of ads sold on their networks, with metadata about the audience demographics and how the ad was targeted. See: Paddy Leersen, Platform ad archives in Article 30 DSA, 25 May 2021, URL: https://dsa-observatory.eu/2021/05/25/platform-ad-archives-in-article-30-dsa/.

[21] European Commission, Questions and Answers: Digital Services Act. "*Once designated by the Commission, providers of very large platforms and very large online search engines have four months to comply with the DSA. Designation by the Commission takes place on the basis of user numbers reported by these services providers, which service providers will have three months after entry into force of the DSA to provide.*" URL: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348

If training data is used: Is the data representative of the context in which the algorithm will be deployed?

Facebook regards the exact inner workings of its systems a trade secret and does not publish information on to what level the data are representative for the context where the algorithms are used, neither does Facebook give access to investigate this.

4.1.4 Security and archiving

Are the data sufficiently secure? Make a distinction here between the input data and the output data.

Facebook does not publish details on its security management nor does Facebook publish any information on security audits and security certifications. See the DPIA for pseudonymisation.

Is there access control over the data? Make a distinction here between the input data and the output data.

Facebook regards the exact inner workings of its systems a trade secret and does not publish details on the access control of the input data used. Facebooks privacy policy states: *"We share the information that we collect globally, both internally across our offices and data centres, and externally with our partners, vendors, service providers and third parties."*[22] Facebook does describe what data it shares with some partners such as advertisers, but in other cases, Facebook only explains 'we share information with...' Facebook does not specify what access controls are in place while granting access to third parties.

How are relevant rules on archiving observed, as laid down in the Archives Act ('Archiefwet')?

Facebook has no facilities in place to comply with the Archives Act. Government organisations are themselves responsible to retain copies of their communications as long as required based on the specific Archive rules for their organisation.[23] See chapter 10 of the DPIA for general information on the retention periods applied by Facebook.

4.2 Part B Algorithm - throughput

4.2.1 Type of Algorithm

Type of algorithm: what kind of algorithm is used or going to be used? How does it work? Distinction between: A non-self-learning algorithm, in which the human specifies the rules that the computer must follow and a self-learning algorithm, in which the machine itself learns about the patterns in the data (machine learning)

See also the answer to 4.1.1. Facebook regards the exact inner workings of its systems a trade secret and does not publish detailed information on the deployed algorithms.

Why is this type of algorithm chosen?

Why is this type of algorithm best suited to achieve the purposes formulated in question 3.2 (Purposes) achieved?

What alternatives are there and why are they less appropriate or useful?

[22] Meta Privacy Center, Privacy Policy, How do we share information with Partners, vendors, service providers and third parties? URL:, https://www.facebook.com/privacy/policy/?section_id=4-HowDoWeShare .

[23] As specified in a Basis Selectie Document for each ministry.

Because it is unknown what algorithms are chosen by Facebook, these questions can not be answered.

4.2.2 Ownership and control

If the algorithm has been developed by an external party: have clear agreements been made about made about ownership and management of the algorithm? What are those agreements?

The algorithms are under ownership and control of Facebook. Government organisations do not have control over updates Facebook makes and is not able to explain how the algorithms work.

4.2.3 Accuracy of the algorithm

What is the accuracy of the algorithm, and on the basis of which evaluation criteria is this accuracy determined?

Is the degree of accuracy (question 2B.3.1) acceptable for the way the algorithm will be deployed?

Facebook does not publish any data on the accuracy of its algorithms therefore these two questions can not be answered.

How is the algorithm tested?

Facebook does not publish any information about its testing processes.

What measures can be taken to counteract the risks of replication or even amplification of biases (e.g. different sampling strategy, feature modification, ...)?

What assumptions underlie the selection and weighting of indicators? Are those assumptions justified? Why are they/aren't they?

Because it is unknown what algorithms Facebook deploys, and in what way, these questions can not be answered.

How often/worse is the algorithm wrong? (e.g., in terms of number of false positives, false negatives, R-squared, ...)

Facebook does not publish any data on the accuracy of its algorithms.

This will change under the DSA. Both the Digital Services Coordinator of establishment or the commission and vetted researchers will have a right to access specific data, including data related to algorithms. This includes the data necessary to assess the risks and possible harms brought about by the platform's or search engine's systems, "*data on the accuracy, functioning and testing of algorithmic systems for content moderation, recommender systems or advertising systems*, *including, where appropriate training data and algorithms*."[24]

4.2.4 Transparency and explainability

Is it clear what the algorithm does, how it does this, and on the basis of what (what data) it does this? Explain.

The algorithms take care of the personalisation, but only Facebook knows how they do so and based on what data. Article 29 of the DSA creates a right to refuse personalised profiling. This article stipulates: "*In addition to the requirements set out in Article 24a,* ***providers of very***

[24] See Recital 64 of the DSA, URL: https://digital-europe-website-v1.s3.fr-par.scw.cloud/uploads/2021/03/FINAL-DSA-Paper-March-2021-1.pdf .

large online platforms that use recommender systems shall provide at least one option *for each of their recommender systems which is* ***not based on profiling****, within the meaning of Article 4, point (4), of Regulation (EU) 2016/679*." If Facebook effects this option, including the use of tracking cookies, and offers Page visitors a fair and clear choice to select the non-profiled option, Facebook can solve the problem of a lack of transparency and explainability.

To which individuals and groups within and outside of the own organisation is the operation of the algorithm made transparent and how is this done?

Only Facebook has insight in the algorithms. The recent Californian class action court case (provisionally settled without judgment in August 2022[25]) suggests that because of the complexity of the inner workings of Facebooks systems even for employees of Facebook it is hard to understand the workings of the algorithms. This was recently confirmed in a hearing of Facebook engineers organised by the Californian Court as part of the settlement.[26]

For which target groups should the algorithm be explainable?

1 Users with a Facebook account, visiting a government Facebook page.
2 Users without a Facebook account, visiting a government Facebook page.
3 The government organisation publishing the Facebook page, both from a legal compliance perspective, and from an employer perspective (responsibility for the admins maintaining the government Page).

Can the operation of the algorithm for the target groups identified in the previous question be explained in a sufficiently comprehensible way?

No, currently not. This might change once the DSA rules for Very Large Platforms enter into effect.

5. Implementation, use and monitoring – outputs

An algorithm as such does not create unwanted effects. These effects are always caused by implementation, deployment or application of the algorithm, by the context in which the algorithm is deployed, and by the decisions and actions that are linked to the output of the algorithm.

This part is therefore about the implementation and use of the algorithm in question, that is, about the (handling of) the "output" of the algorithm.

5.1 Decisions based on output algorithm

What happens to the results of the algorithm? What decisions are based on them?

Facebook uses results of the algorithms for the personalisation of:

1. Postings shown in the News Feed from people or organisations the user has chosen to follow, for example the position of the government Page in the News Feed.
2. Postings shown in the News Feed based on Facebook's algorithmic recommendations, for example to who the government Page is recommended or what other pages are recommended on the government Page.
3. Advertisements shown as 'sponsored content' on the government Page.

[25] The Register, Facebook settles Cambridge Analytica class action for undisclosed amount, 29 August 2022, URL: https://www.theregister.com/2022/08/29/facebook_settles_cambridge_analytica_lawsuit/
[26] The Intercept, Facebook Engineers: We Have No Idea Where We Keep All Your Personal Data, 7 September 2022, URL: https://theintercept.com/2022/09/07/facebook-personal-data-no-accountability/

4. Advertisements shown as ‘sponsored posts’ on the government Page.
5. Advertisements shown as ‘related pages’ when visiting the government Page.

The government organisation does not base any decisions on the results of these algorithms.

5.2 The role of humans in the decision

What role do people play in making decisions based on the output of the algorithm ('human in the loop') and how are they enabled to play that role?

There are no ‘humans in the loop’ when Facebook determines the personalisation.

5.2.1 Is sufficient qualified personnel available, now and in the future, to manage, review and adapt and modify the algorithm if desired/needed?

Facebook has sufficient qualified personnel available to modify the algorithms. According to its latest annual financial report, Meta employs 71,970 people.[27]

5.3 Effects of the algorithm

What will be the effects of the deployment of the algorithm for citizens and how will the ‘human dimension’ be taken into account when making decisions based on the output of the algorithm?

There is too much unknown about the deployment of the algorithms by Facebook to reliably assess the effects of the deployment of the algorithm for citizens. Facebook does not publish claims that it takes the ‘human dimension’ into account when making decisions based on the output of the algorithm.

What are the risks of stigmatising, discriminating or otherwise harmful or adverse effects on citizens and how will these be mitigated?

There is too much unknown about the deployment of the algorithms by Facebook to reliably assess the risks of stigmatising, discriminating or otherwise harmful or adverse effects on citizens. Facebook does not publish information on how it mitigates these effects.

This will change in the DSA. See for example Recital 56 about one of the four risks that must be mitigated: “*actual or foreseeable impact of the service on the exercise of fundamental rights, as protected by the Charter, including but not limited to human dignity, freedom of expression and information, including media freedom and pluralism, the right to private life, data protection,* ***the right to non-discrimination****, the rights of the child and consumer protection. Such risks may arise, for example,* ***in relation to the design of the algorithmic systems used by the very large online platform*** *or by very large online search engine or the misuse of their service through the submission of abusive notices or other methods for silencing speech or hampering competition.*”

How will the expected effects contribute to solving the problem that caused the development/deployment of the algorithm (see question 1.1) and achieving the purposes set for it (see question 1.2)?

There is too much unknown about the deployment of the algorithms by Facebook to reliably assess their contribution to solving the initial problem and achieving the purposes.

How do the expected effects relate to the values being served (see question 1.3)? What risks are there that certain values will be under pressure and how will this be dealt with?

[27] https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_FB_2021.pdf

There is too much unknown about the deployment of the algorithms by Facebook to reliably assess their effects on the values being served, to assess the risks of values being under pressure. Because these are unknown, it is impossible to say they are dealt with.

5.4 Procedures

By what procedures will decisions be based on the algorithm?

The exact procedure Facebook uses to take decisions on basis of its algorithms is unknown.

How are different relevant actors (administrative, political and citizens) involved in the decision-making process?

Government organisations and Page visitors are not involved in the decision-making process.

How is it ensured that the requirements of good and proper administration and - where necessary - legal protection are met in these procedures? Do citizens have an effective possibility to lodge a complaint or objection? If so, in what way?

Government organisations do not have control over the procedures, and no means to ensure the requirements of good and proper administration are met. Facebook did not provide the researchers with any meaningful information about the algorithms and decision making in reaction to the data subject access requests performed for the DPIA (Section 2.5). Citizens have therefore no effective possibility to understand the decisions made and no effective possibility to lodge a complaint.

5.5 Context

Time/period: when will the algorithm be deployed? How long is the period that it will be deployed?

Many Dutch government organisations already use Facebook Pages. No end date has been set.

Place: where will the algorithm be deployed? Is that in a particular geographical area, is that with a particular group of people or files?

The algorithms will be deployed whenever a person visits a government Facebook page.

Can the algorithm still be used if context factors change or if the algorithm is used in a different context than that for which it was developed?

There is no documentation about the context for which the algorithms were developed. Facebook regards the details about the algorithm as trade secret. Therefore it is impossible to judge what is a context change and if the algorithm still can be used.

5.6 Communication

How open can you be about the working of the algorithm in the light of the objectives and context of its deployment?

Facebook regards the working of the algorithms a trade secret. Government organisations with Facebook Pages do not have information about the working of the algorithms. It is not possible to be open about the working of the algorithms.

How do you intend to communicate about the use of the algorithm?

The updated Facebook Privacy Policy[28] does not mention the deployment of algorithms, but it does mention several times there is automated processing without any further details on this automated processing. Government organisations with Facebook Pages cannot meaningfully communicate about the use of the algorithm.

Is the output of the algorithm visualised for example in a table, graph or dashboard? If so: is the form of visualisation or display a correct representation of the output of the algorithm? Is the visualisation easy to read for different for different user groups?

The output of the algorithms is not visualised.

5.7 Evaluation, auditing and assurance

Are there good tools for evaluation auditing and assurance of the algorithm?

Government organisations with Facebook Pages do not have good tools for evaluation, auditing and assurance of the algorithms.

Are there sufficient possibilities to account for the algorithm?

Government organisations with Facebook Pages do not have sufficient possibilities to account for the algorithms.

What possibilities are there for auditors and supervisors to attach (formal) consequences to the use of an algorithm by the government (e.g. feedback of findings, making recommendations, budgetary consequences, etc.)?

Government organisations with Facebook Pages do not have a meaningful way to follow-up feedback, findings or recommendations. Budgetary consequences are possible in case the government organisation receives a punitive sanction.

This will change once the DSA rules enter into effect for Very Large Platforms such as Facebook. National Digital Services Coordinators that collaborate in a Board, as well as the European Commission will have enforcement powers. See Chapter IV of the DSA, *Implementation, cooperation, sanctions and enforcement*. Additionally, Section 4 (additional obligations for providers of very large online platforms and very large online search engines to manage systemic risks, art. 25 to 33) imposes many reporting, risk assessment and independent audit obligations on providers of Very Large Platforms.

6. Human rights

6.1 Introduction

The first three parts of the IAMA model contain questions and concerns that are relevant to all algorithms and for which serious and proper discussion can help ensure that algorithms are deployed in a careful, well-considered and well-embedded manner. In view of the great importance of the protection of human rights and the special risks that may exist for the impairment of these human rights through the use of algorithms, it is important to pay separate attention to this topic.

This part of the IAMA model therefore contains a human rights roadmap', which has a twofold purpose:

[28] Facebook Privacy Policy (entered into effect on 26 July 2022), URL: https://www.facebook.com/privacy/policy

1 It serves as a tool to identify whether the algorithm to be deployed will affect human rights;
2 If so, it enables a structured discussion about the question of whether there are possibilities to prevent or mitigate this human rights violation, and whether there are grounds on which the (mitigated or not) human rights violation should be found acceptable after all.

Several steps of the IAMA model are also relevant when answering the question which public values can be served and affected by an algorithm.

For each human right impacted by the algorithms, the IAMA model asks seven questions:

1 **Human right**; Which (aspect of) a human right is affected?
2 **Specific legislation and directives**; Is there any specific legislation applicable to this human right? If so, is it complied with?
3 **Weighting**; What aspect of the human right is affected and is the expected interference far-reaching, limited or something in between? What is the corresponding colour code (red, orange, green)?
4 **Purposes**; What purposes are to be achieved? See answer to question 1.2.
5 **Effectiveness/appropriateness**; Is the algorithm to be used an efficient, suitable and effective means of achieving the goals set?
6 **Necessity/subsidiarity**; Is deployment of this specific algorithm necessary, i.e. are there no other means or mitigating measures available to do so?
7 **Assessment/proportionality in the strict sense**; Are the objectives, on balance, weighty enough to justify the violation of the infringement of the human right?

This HRIA has shown so far that many aspects of the algorithms deployed are unknown and that it is not possible to assess their impact. Therefore it is impossible to answer question 2 to 7 of this model. Therefore this HRIA takes a different approach and answers for each aspect of the human rights the following answers:

1 **Potential impact**; Is this aspect of human rights potentially impacted when the government uses Facebook Pages? If so, in which way?
2 **How to assess**; How can the potential impact on this aspect of human rights be assessed?

2 Analysis per aspect of human rights

.1 *Rights appertaining to the person*

Cluster	Examples	**Potential impact**	**How to assess?**
Personal identity / personality rights / personal autonomy	- Right to development - Freedom to determine one's own actions - Freedom to shape one’s own appearance (Freedom of dress) - Free choice of profession, choice of education, choice of training etc. (Freedom to choose an occupation) - Respect for one's own identity (gender identity / sexual identity etc.)	These processes are performed by identification (see analytical framework). Identification is only possible when different groups and minorities people can identify themselves with, are visible and represented. The personalisation of Facebook can potentially both increase and decrease the diversity in the representation, resulting in enhancing or diminishing this right. Facebooks policy that requires users to register their account under their legal name, results in users being engaged according to properties that may be apparent from that name, like gender or inheritance and not in the way they identify themselves or want to engage with other people themselves. This may impact the way visitors react to government Pages.	Test in a laboratory setting if following a government Page results in differences in the representation of minorities. Research on how interactions on Facebook change if people post under pseudonym.
	- Reproductive rights - Right to know one's own parentage (Right to know one’s own biological family) - Name rights (Right to respect for one’s name) - Freedom of contract	No direct impact.	

Cluster	Examples	**Potential impact**	**How to assess?**
Social identity / relational privacy rights / relational autonomy	- Right to respect for family relations/family life (Right to respect for private and family life) - Right to marry - Right to found a family - Right to enter into sexual relations - Right to enter into professional/business relationships - Right to access employment/profession (Right to engage in work) - Right to enter a country / rights of residence - Right to education	No direct impact.	
Physical and mental integrity	- Freedom of conscience / freedom of thought	This right depends on a social environment that is open to and supportive for different world views and different approaches to ethics. A social network potentially offers a platform for diversity in views but the personalisation can steer people into their own bubble. It is unclear to what extent visiting or following a Page on Facebook contributes to such a bubble.	Test in a laboratory setting if following a government page results in an increase or decrease of different views represented in the personalisation.
	- Right to life - Prohibition of torture / inhuman or degrading treatment and punishment - Prohibition of searching: body search (Right to be free from illegal body searches)	Can be impacted in extraterritorial cases.	DTIA covering all possible data transfers and extra territorial laws.

Cluster	Examples	**Potential impact**	**How to assess?**
	- Prohibition of refoulement - Requirement of consent in medical treatment and research - Right of access to health care - Respect for legal capacity - Right to voluntary termination of life - Right to abortion - Prohibition of (modern) slavery/servitude/forced labor/human trafficking/exploitation	No direct impact.	
Data protection / informational privacy rights	- Protection against unlawful/negligent data processing - Right of access - Right to rectification - Right to be forgotten	See DPIA.	
Communication rights	- Letter secrecy (secrecy of correspondence) - Protection against eavesdropping/tapping/interception - Prohibition of unlawful transfer of communications data - Confidentiality of communications with an attorney, physician, etc.	See DPIA.	
Spatial privacy rights	- Freedom of movement - Habeas corpus rights (prohibition of deprivation of liberty, house arrest, etc.) - Free choice of residence (Freedom of residence)	Can be impacted in extraterritorial cases.	DTIA covering all possible data transfers and extra territorial laws.
	- Prohibition of unlawful tracking of persons (GPS tracker)	See DPIA.	

Cluster	Examples	**Potential impact**	**How to assess?**
	- Free movement rights (EU law) (Freedom of movement) - Right to leave the country - Prohibition of unlawful camera surveillance	No direct impact.	
Property bound privacy rights	- Protection from search of clothing/bags/laptop/computer etc.	Can be impacted in extraterritorial cases.	DTIA covering all possible data transfers and extra territorial laws.
	- Right to respect for home (protection from raids/searches) - Free disposal of property (Right to property) - Protection against expropriation - Intellectual property rights	No direct impact.	
Reputation Rights (The right to protection of reputation)	- Prohibition of criminal defamation / libel / slander - Protection of honour and good name	Though a social network can have a major impact on this, the act of visiting or following a Facebook page does not seem to impact this.	
Healthy living environment (Environmental protection)	- Right to sustainable development - Right to environmental protection - Protection from emissions of harmful substances - Right to clean drinking water - Right to sanitation (sewage) - Right to access energy	No direct impact.	
Social and economic rights	- Right to a minimum level of subsistence - Right to social security and assistance - Access to education (Right to education)	No direct impact.	

.2 *Freedoms*

Cluster	Examples	Potential impact	How to assess?
Freedom of expression	- Press freedom/journalistic freedom - Artistic freedom - Freedom of science/academic freedom - Freedom to choose means of expression (oral/written etc.) - Whistleblowing (Freedom to speak up) - Journalistic source protection/disclosure rights	A social network can be a platform for freedom of speech. At the same time its own policies can limit freedom of speech. Also a bias in the personalisation may push users in a certain direction. The visibility of directly identifiable users, both for other users and for authorities may also result in a *chilling* effect. Messages may also be coloured by what users expect they have to do to get a high ranking. Visiting or following a Facebook Page may have an impact on the personalisation and consequently, on the right to freedom of expression.	Monitor and document cases where Facebooks own policy collides with freedom of speech rights. Test in a laboratory setting if following a government page results in an increase or decrease of different views represented in the personalisation. Behavioural research on possible chilling effects.
Freedom to receive information	- Passive information-gathering (right of access to existing information)	By publishing information on a Facebook Page where people are already active, the Dutch government actively supports this right. On the other hand, because Facebook pages are only partially accessible without a Facebook account, free access is diminished by using Facebook Pages.	
	- Active information-gathering (right of access to public information / open government)	No direct impact as long as government organisations make all information also available on public webpages, outside of the Facebook platform.	
	- Duty to provide pluralistic information	The personalisation of Facebook may result in more uniform information.	Access for vetted researchers to actual data processed by Facebook

Cluster	Examples	Potential impact	How to assess?
			relating to popular government Pages, to investigate if the following of a government Page by a user, or by their friends, results in an increase or decrease of different views represented in the personalisation. Additionally, researchers must be able to perform A/B testing in an isolated lab, with model accounts. Currently, Facebook prohibits the use of test accounts.
	- Right of free access to the internet (Right to internet access)	No direct impact.	
Freedom of religion	- Freedom to have or not to have a religion - Freedom of religious expression (symbols, rituals) - Freedom to congregate with other believers - Freedoms of denominations/religious communities - Separation of state and religion (religious neutrality of the state) - Respect for religious/philosophical beliefs in education	No direct impact.	
Freedom of assembly and protest	- Freedom of assembly - Assemblies, protest marches, etc. - Free choice of subject, time, place and means - Protection from 'hostile audiences'	Social networks play an important role in organizing protests. Posting on a Facebook page makes it easier to organize demonstrations about that topic. The personalisation may push users towards radicalization and more radical protests.	Access for vetted researchers to actual data processed by Facebook relating to popular government Pages, to investigate if following a government Page results in an increase or decrease of different views represented in the personalisation. Additionally, researchers must be able

Cluster	Examples	Potential impact	How to assess?
			to perform A/B testing in an isolated lab, with model accounts. Currently, Facebook prohibits the use of test accounts.
Freedom of association	- Freedom to be or not to be a member of an association - Internal association freedom (own choice of members, activities) - Freedom of political parties (Freedom of assembly and of association in political matters) - Trade union freedom (Freedom of assembly and of association in trade union matters) - Collective action rights - Right to strike	No direct impact.	
Political rights/freedoms	- Right to periodic organization of free and secret elections (Right to free elections) - Right to vote and to stand as a candidate for election - Right to petition	No direct impact.	

.3 *Equality rights*

Cluster	Examples	Potential impact	How to assess?
Equality before the law	- Equal application of general legislation to all who fall within its reach - Prohibition of arbitrariness - Requirement of consistency - Principle of legal certainty	No direct impact.	

Cluster	Examples	Potential impact	How to assess?
Prohibition of direct discrimination on certain grounds	- Decisions or rules must not be overwhelmingly motivated by or based on protected personal characteristics	No direct impact.	
Prohibition of indirect discrimination on certain grounds	- Decisions or rules must not disproportionately disadvantage persons belonging to groups with protected personal characteristics	The Page that is visited or followed or the messages that are read on the Page might correlate with protected properties. When the Pages that are visited are used to determine the personalisation, this can result in indirect discrimination.	Access for vetted researchers to actual data processed by Facebook relating to popular government Pages, to investigate if viewing certain information on a government page results in differences while the algorithms deduce properties or select groups, e.g. for advertising. Additionally, researchers must be able to perform A/B testing in an isolated lab, with model accounts. Currently, Facebook prohibits the use of test accounts. Access for vetted researchers to actual data processed by Facebook to analyse if there is a correlation between non-protected properties processed by Facebook and protected properties.
Prohibition of discriminatorily motivated action	- Prohibition of racist/xenophobic etc. motivated action (e.g. assault) - Prohibition of ordering discrimination	No direct impact.	
Right to material distinction / customization	- Duty to take into account differences between people and groups	No direct impact.	

Cluster	Examples	Potential impact	How to assess?
Right to reasonable accommodation / affirmative action	- Right to facilities for people with disabilities that enable them to participate equally in society	Facebook invests in accessibility of their Pages. Facebook has no accessibility certification. Communication with Facebook pages is inaccessible to people who no to limited access to the internet or technical abilities.	Test for WCAG 2.1 compliance.
	- Right to compensatory measures for past structural social inequality	It is not possible to adjust the algorithms of Facebook to compensate for inequalities the Dutch government decides to compensate for. It is possible that the personalisation reinforces such inequalities for example by ranking down opportunities for certain groups when they are published on a government Page.	
Prohibition of profiling	- Prohibition of the creation of categories or profiles based on protected personal characteristics, which then form the basis for decision-making or policy (The right not to be subject to a decision based solely on profiling)	Facebook asks users to provide protected properties like age and gender. It is not clear to what level these protected properties influence the personalisation of Facebook.	Analyse if there is a correlation between non-protected properties processed by Facebook and protected properties, this must be done with access to Facebooks databases. Test in a laboratory setting if these protected properties and correlating properties influence the personalisation.
Prohibition of segregation	- Prohibition of spatial or other forms of separation of groups that do receive similar treatment in the process	Potentially the personalisation can result in groups that see comparable personalisation within the group, but that differs from the personalisation of other groups because they are not seeing anything from other groups. This can effectively result in a	Make a network analysis of the suggestions given to a large group of Facebook users. Test in a laboratory setup the impact of following or viewing a government Page on the networks. To perform these tests, researchers must be able

Cluster	Examples	Potential impact	How to assess?
		segregation of groups. To what extent this happens and the influence of visiting of following a government Page on this is unknown.	to perform A/B testing with model accounts. Currently, Facebook prohibits the use of test accounts.

.4 Procedural rights

Cluster	Examples	Potential impact	How to assess?
Right to proper administration (pre-trial stage)	- Right to transparency and information	By publishing information on a Facebook Page, Dutch government organisations make themselves transparent to the users of Facebook. At the same time the personalisation can result in people not being notified of new information, diminishing transparency.	Analyse what groups do see and what groups don't see suggestions for the government page because of the personalisation.
	- Participation and defence rights (e.g., hearing rights) - Right to careful decision-making - Obligations to state reasons (Right to a reasoned decision) - Prohibition of arbitrariness - Prohibition of abusive conduct/détournement de pouvoir	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and followed government Pages, might be used in prosecutions that violate these rights.	DTIA covering all possible data transfers and extra territorial laws.
Right to an effective remedy and access to justice (trial phase)	- Power to provide effective redress	It is unclear to what extent Facebook is able to cooperate when its involvement is needed to perform this human right.	Monitor and document cases where Facebook does or does not correct unjust decisions.
	- Prohibition of high thresholds (court fees, assistance of a counsel, immunities, time limits)	The reaction to the data subject access request is not complete (see DPIA), so information possibly needed for a fair trial needs to be claimed with	

Cluster	Examples	Potential impact	How to assess?
		a different procedure and thus this human right might be violated.	
	- Right to access to a state court - Right to a substantive hearing - 'Full jurisdiction' (entire case must be able to be adjudicated) - Ius de non evocandi (right not to be deprived of the justice due to someone in a particular dispute) (Right to access justice) - Right to legal aid - Right to effective enforcement of a court judgment	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and followed government Pages, might be used in prosecutions that violate these rights.	DTIA covering all possible data transfers and extra territorial laws.
Right to an independent and impartial judge	- Personal independence (e.g., tenure) - Business/functional independence (protection from outside pressures) - Institutional independence - Subjective impartiality (no involvement with any of the parties) - Objective impartiality (no legitimate doubt regarding a biased judgment)	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and followed government Pages, might be used in prosecutions that violate these rights.	DTIA covering all possible data transfers and extra territorial laws.
Right to a decision within a reasonable time	- Entire procedure (including preliminary procedure) should not take too long - Right to means to speed up the procedure - Right to compensation if the procedure takes too long	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and followed government Pages, might be used in prosecutions that violate these rights.	DTIA covering all possible data transfers and extra territorial laws.
Right to a fair trial	- Adversarial procedure - Equality of arms (preparation time, access to files/documents) - Claim and counterclaim	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and	DTIA covering all possible data transfers and extra territorial laws.

Cluster	Examples	Potential impact	How to assess?
	- Balanced and fair rules on evidence - Obligation of the judge to state reasons - Equal opportunities for parties to hear witnesses / experts - Publicity at the hearing/judgment - Legal certainty - Rights of non-disclosure (confidentiality)	followed government Pages, might be used in prosecutions that violate these rights.	
Right to a fair trial: criminal justice safeguards	- Presumption of innocence - Right to remain silent/right not to cooperate - Prohibition of incitement - Ne bis in idem (right not to be tried or punished twice) - Right to assistance by lawyer (legal assistance) - Right to assistance by an interpreter	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and followed government Pages, might be used in prosecutions that violate these rights.	DTIA covering all possible data transfers and extra territorial laws.
Criminal legality requirement (no punishment without law)	- Prohibition of retroactive application of criminal laws - Lex mitior principle	In case of an extra-territorial prosecution (see analytical framework) the data collected by Facebook, including visited and followed government Pages, might be used in prosecutions that violate these rights.	DTIA covering all possible data transfers and extra territorial laws.

Conclusion

The purpose for Dutch government organisations to use Facebook Pages is to increase the reach of their communication. When a person visits a Dutch government Page, Facebook uses the information about this visit, and possible interactions with the content on the Page to show a personalised *News Feed* and personalised advertisements. Government organisations cannot control this personalisation. Facebook uses an automated system, with several algorithms, probably both self-learning and non-self-learning, to select and rank the information shown in each user's personal News Feed and advertisements. Facebook does not provide any detailed information on the development and inner workings of these algorithms, neither in public communication nor in response to the Data Subject Access Requests filed for the DPIA. This HRIA concludes that many questions about the algorithms cannot be answered because Facebook does not provide the required information.

Facebook is not transparent about:

1. The types of algorithms deployed.
2. The data sources used while composing and/or training the algorithms, the quality of those data, possible biases within those data and how representative those data are.
3. How both the input and the output data are secured and how long the data are retained.
4. The design choices made while creating the algorithms.
5. The accuracy of the algorithms.
6. The testing procedures for the algorithms.

Nor does Facebook offer an interface for government organisations to structurally or incidentally overrule the decisions made by the algorithms. The lack of an interface makes it impossible for government organisations to correct decisions or to 'put a human in the loop'. Facebook does not provide any tools either for government organisations to evaluate or audit the algorithms. Without such tools, government organisations cannot explain to their visitors how the algorithms are deployed. They also cannot ensure the personalisation follows affirmative actions of government organisations or provide evidence when a decision by the algorithms is contested in court.

This lack of transparency and lack of auditing tools precludes government organisations from judging the effects, positive or harmful, of the algorithms Facebooks deploys. Government organisations cannot assess if any mitigating actions are needed to counter harmful effects and if so, what actions are needed.

When a government organisation uses Facebook Pages, the personalisation of Facebook potentially has an impact on several human rights. Causes for these potential impacts are the possibility that the algorithms steer towards less diversity or representation of minorities, the possibility that the algorithms are pushing the users towards certain actions, opinions or lines of thought, are pushing users to extremes, or because the personalisation might be discriminating and might be segregating groups. These effects potentially impact the following human rights:

1. Personal identity / personality rights / personal autonomy
2. Freedom of conscience / freedom of thought
3. Freedom of expression
4. Freedom to receive information
5. Freedom of assembly and protest
6. Prohibition of indirect discrimination on certain grounds
7. Prohibition of profiling
8. Prohibition of segregation
9. Right to transparency and information

Government authorities in third countries may obtain data about visitors of Dutch government Pages. These authorities may compel disclosure, use OSINT or use ADINT to trace visits or visitors of government Facebook Pages. This may even, as ultimate consequence, result in intimidation, (cyber)attacks or unfair trials, or in human right violations when such visitors travel to third countries. Human rights that are potentially impacted by these extraterritorial effects are:

1. Physical and mental integrity
2. Spatial privacy rights
3. Protection from search of clothing/bags/laptop/computer etc.
4. Right to proper administration (pre-trial stage)
5. Right to an effective remedy and access to justice (trial phase)
6. Right to an independent and impartial judge
7. Right to a decision within a reasonable time
8. Right to a fair trial
9. Right to a fair trial: criminal justice safeguards
10. Criminal legality requirement (no punishment without law)

These potential impacts on human rights are separate from the data protection risks assessed in the DPIA.

The observations in the DPIA show that bias by one sided suggestions is present and that bias by feedback loops and bias by optimalisation criteria are very likely to be present. This optimalisation bias is possibly caused by the use of the amount of likes as optimalisation criterium. This effect may reinforce itself because the highest suggested posts are the most likely to receive new clicks and likes.

In sum, this HRIA concludes the use of Facebook Pages by Dutch government organisations has a high potential impact on at least 9 human rights, plus 10 extra rights when data about visits to a Page are used by government authorities in third countries. Due to Facebook's lack of transparency and lack of assessment tools, it is impossible to assess the impact in reality. Because government organisations have no means to assess, control, correct or explain the algorithms used, the use of government Facebook Pages must be qualified as a high risk for human rights.

This situation will hopefully improve early in 2023, when specific rules from the Digital Services Act for Very Large Platforms such as Facebook enter into force. To properly access the lawfulness, ethical impact and robustness[29] of the algorithms deployed by Facebook for personalisation, it is essential that the European Commission as designated supervisory authority for Facebook, as well as vetted researchers, will obtain access to the development process of Facebook and the choices already made, to the data stored by Facebook about government Page visits, not limited to the data used to produce Page Insights, as well access to a controlled environment that enables observing the behaviour of the algorithms.

[29] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai