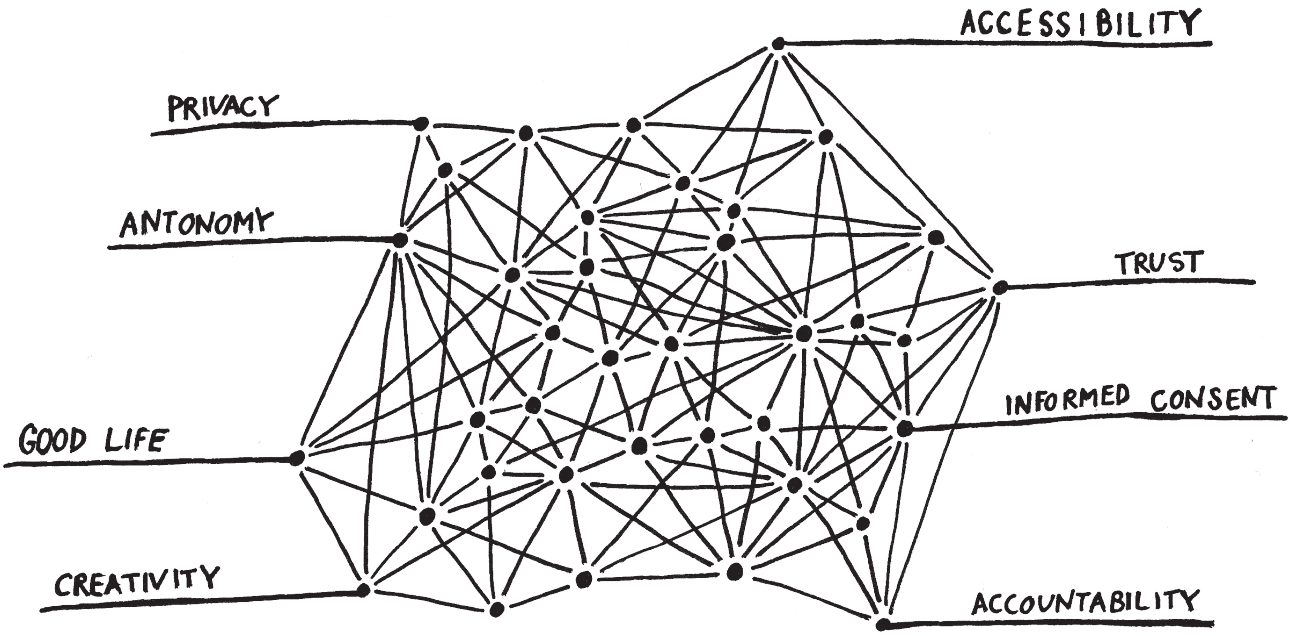
D *ata*

E *thics*

D *ecision*

A *id*



Inventarisation of **ethical issues**
emerging from **municipal data projects**

Disclaimer

DEDA version 1.0,
released March 2017,
check for updates:

www.dataschool.nl/deda

Neither the Utrecht Dataschool, the University of Utrecht nor the Municipality of Utrecht are liable for (possible) damage in connection with, or arising from, the use of material or content of DEDA.

Utrecht, 2017.

DEDA version 1.0.

While DEDA is used we learn how to improve it. Those changes will be implemented in future versions. If you have used DEDA and you think it can be approved, please feel free to share your thought with us. Send an e-mail to: **info@dataschool.nl** - We appreciate your feedback.”

Concept and development
Aline Franzke, MA

Supervision:

Mirko Tobias Schäfer, Ph.D.

Commissioned by :
Gemeente Utrecht

Design:
Redmar Fransen

Illustration:
Sammy Hemerik

Table of content

Intro

Why now?

‘Big data’ and novel practices of analysis promise great benefits for public management. Such promises reach from facilitating smart cities over advancing economic prosperity to empowering citizens. However, these opportunities also bear difficulties that are easy to neglect but might eventually turn good intentions into bad results, constitute liabilities and violate conducts of good governance. Citizens are increasingly alarmed about the way corporations or governments use data. In response, several practices have become regulated and laws have been adapted. The aggravated fees for privacy violations are an example of the EU’s efforts to enforce a responsible use of personal information. But next to privacy, a number of other issues emerge from data projects: data sets taken out of context or of questionable origin; bias in data sets, models and algorithms; the questions of conflict of interests of commercial suppliers and public institutions; and the social impact of data-driven policies and how to evaluate them critically are just a few fields where the law does not always apply or provides a clear guideline for action. These gray areas informed by commonly shared values and social responsibility can be addressed through guidelines for ethical decision making.

What is DEDA?

DEDA, acronym for data ethics decision aid, is a tool to detect ethical issues, develop sensibility for values that might be affected by a data project, and to document the ethical decision making process. DEDA was developed in close cooperation with municipal data analysts. It supports you in responsibly using data, models and algorithms.

Purpose

This document serves two objectives. Firstly, it provides additional information, explanations and examples related to concepts addressed in the DEDA poster. Secondly, it is designed to collect the answers and document the outcome of the reflection process. In this manner this document can be used as a report. The DEDA poster and the DEDA manual correspond and should be used together.

How to use

1 Run through the questions written on the DEDA poster. The questions are structured along a process and address general concerns as well as data-specific concerns. You can use Post-its to freely collect your thoughts on the poster. For each cluster you will find some explanations in this booklet. After the reflection session, fill in the answers in the report at the end of this booklet.

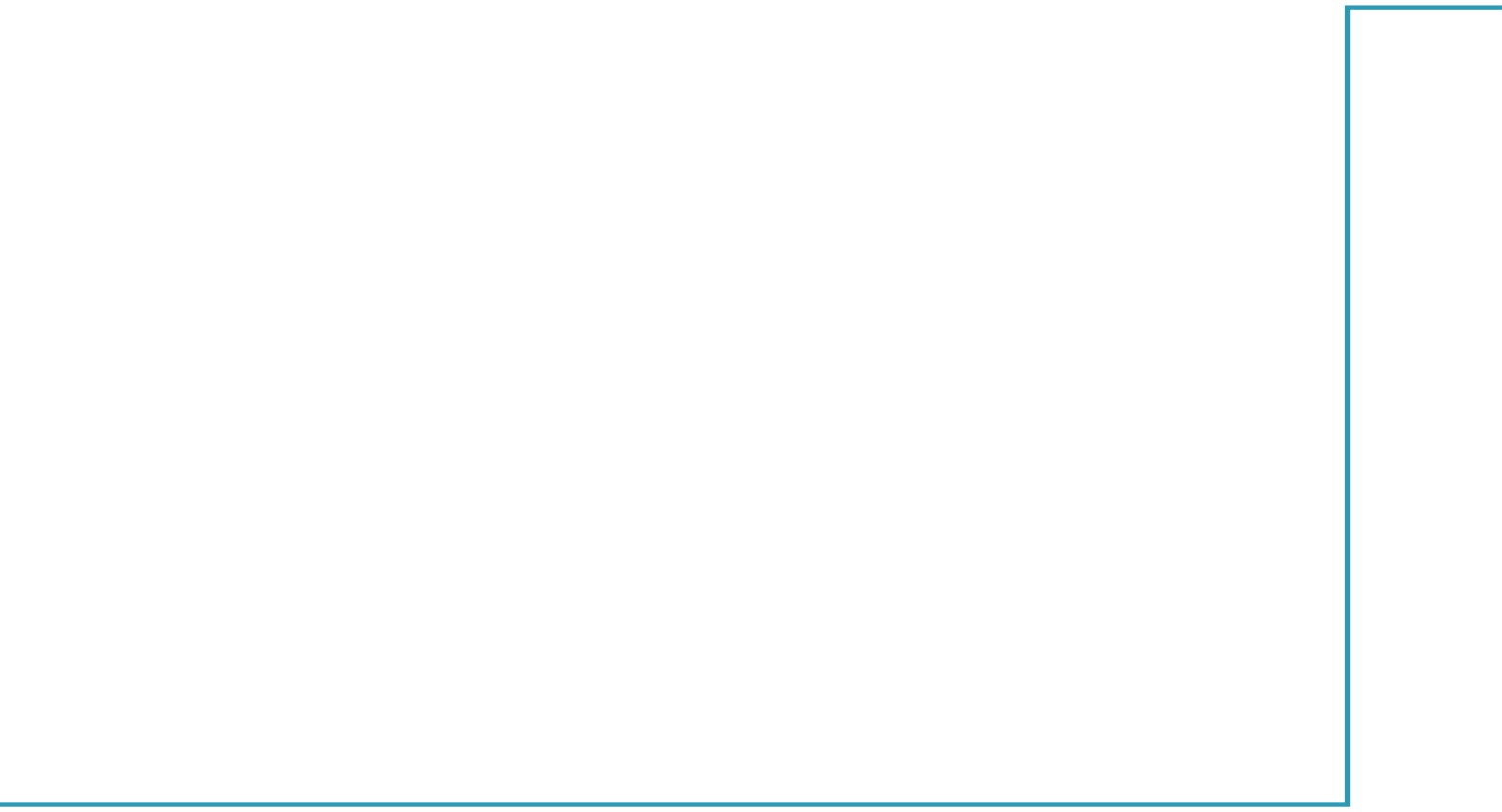
2 Data-specific considerations are marked in blue. In the data-specific section you might skip the clusters not relevant to your current phase of the project. General considerations are marked in green and focus on responsibility, transparency, bias and privacy.

Notice: Questions that can not be answered directly you find an action point at the bottom of the answerpages in the report.

3 The last section on the DEDA poster and in this document provides an ethical decision making process. Based on your process of deliberation concerning the various questions you can now decide whether to move forward with the project and which ethical safeguards you are going to apply to meet your standards of compliance.

Questions

DATA RELATED CONSIDERATIONS



ALGORITHMS

Explanation:

Algorithms can process data and develop insights from information they receive. Using models, they can weigh certain information over other information and produce an output that can be used for decision making: e.g. an algorithm that determines how much parking space is available in which garage; or an algorithm that assesses who is eligible for social welfare and who is not. Algorithms use mathematical models. Models are like opinions about a specific phenomenon. Their opinion is expressed in values and calculations. Algorithms will increasingly inform decision making. It is important to understand how their output has been developed. Algorithms need to be transparent, that means accessible for external experts for scrutiny and verification of their results. Public management institutions must be able to explain how their models and algorithms work. Models and algorithms are also subject of accountability and good governance. It might be unclear who the owner of the algorithm is. Further, the difficulty arises that models and algorithm might not be publicly documented.

Questions

1a **Is there someone in the team who can explain how the algorithm in use works?**

1b **Can you communicate that with the public?**

SOURCE

Explanation:

When data is collected it is important to check the quality of this collected dataset. Frequently the belief “the more the better” is a wide spread credo. However, when data gets collected, the belief ‘the more the better’ is a wide-spread credo. However, this credo might be problematical and should be challenged. The size of bigger datasets does not always increase the quality of what can be reached with this data. Ethical reflections concerning the source of datasets should therefore start with a reflection on what type of data is really needed to collect or to buy.

Questions

2 **Where did your source get the data(sets) from?**

3 **Have you checked the quality of the data(sets)?**

4 **Is there a ‘best before’ date for this specific data set?**

ANONYMIZATION

Explanation:

Why should we bother about Anonymization? When we talk about privacy, anonymization is the necessary means to reach the standard of privacy. Data that are not profoundly anonymized might bring for example the ethical danger of exposing one particularly individual, his or her interests and preferences. This is particularly important because this information can have further interest for companies and third parties that might use this information for behavioural manipulation or personified advertisement. Autonomy of citizens might be in danger due to a lack of anonymization.

Questions

5 Are the data anonymized or pseudo-anonymized?

Explanation question 5
Anonymizing and pseudoanonimiseren are two different ways to edit the technical data to hide the identity of a specific individual. Anonymizing is a process in which the identity of a person is no longer traceable. Pseudoanonimiseren means that privacy sensitive data that is associated with a person, is disengaged from this person. By applying this process, it becomes difficult to identify a specific individual.

6a Have you tested the anonymization?

6b Who is in possession of the encryption key?

VISUALIZATION

Explanation:

In order to visualize data, it has to be cleaned. These decisions, which determine which data are in or out and also determine the choice of visualization style and technique can produce bias or even appear manipulative. Note that one and the same data set visualized with different algorithms can lead to completely different visualizations which suggest different ‘readings’ of the data. Further, not all results of datasets are suitable for visualization.

Questions

7 **Are the data or the produced results suitable for visualization?**

8 **How could this visualization look?**

9 **What would be a different interpretation of this visualization?**

ACCESS

Explanation:

Access can mean different things depending on the context. Here we focus on the question of access to the collected and stored datasets within your organisation. Access is a relevant topic because not every dataset should be freely accessible, to protect confidentiality of citizens. A second common issue with access is that third commercial partners might be interested in datasets, which might cause further ethical challenges that should be investigated with care. The question of who should have access to the data will further be relevant when we reflect about the question of open access for everyone.

Questions

10a Who has access to the dataset?

10b How is the access monitored?

OPEN ACCESS & REUSE OF THE DATASET

Explanation:

Making datasets available for other partners can bring the benefit of participation and transparency. In the same time, sensitive data might become public and commercial partners might misuse data.

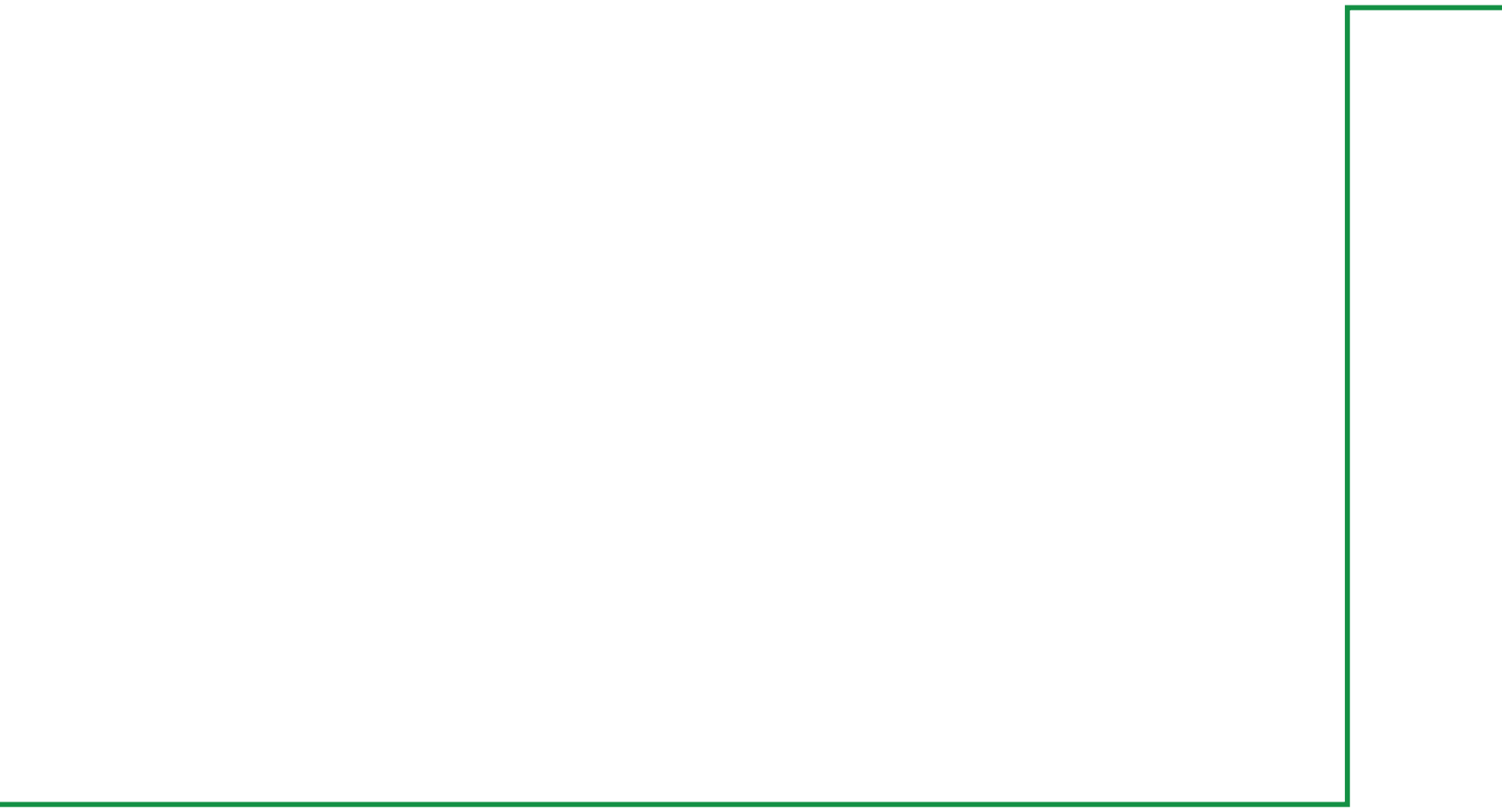
Questions

11a Are parts of the data suitable to be reused?
If yes, which potentials do you see with reusing data?

11b What are possibilities?

12 What dangers do you see with reusing data?

GENERAL CONSIDERATIONS



RESPONSIBILITY

Explanation

Responsibility corresponds in general with the codes of conduct of your specific discipline, your organisation and rules applicable to your specific position. For public servants in the Netherlands this would be the Gedragscode Ambtenaren. Basic values from this code of conduct inform responsible work with data:

- **Good governance**
- **Confidential use of information (e.g. protecting data)**
- **Responsible use of public resources and infrastructures**
- **Conflict of interests**

It must be the general concern of governmental institutions and firms alike to provide just and accountable governance in the best interest of its citizens. Data projects often have an impact on the livelihood of citizens. Keep in mind that political parties, citizens, lawyers or activists might use their rights to inquire about your data projects.

Questions

13 Which laws and regulations are applicable to your project?

14a Can you name a responsible person in your project?

14b Who will take care if something goes wrong?

15 Is there a danger that particular people or groups could be discriminated by your project?

16 Assess who would be suitable partners for your project:

Explanation for question 16

Price is not the main relevant factor; other aspects must be considered as well, such as:

- *What is the origin and quality of the data the external partner promises to provide?*
- *Who will own models, algorithms or data developed during the project?*
- *Can you inspect, access, explain models or algorithms used or developed by the external partner?*
- *Will you become dependent on the external partner?*

17a In case something goes wrong, are there any communication strategies?

17b Who is responsible for preparing those strategies?

TRANSPARENCY/ ACCOUNTABILITY

Explanation

Public management is held accountable by citizens and political parties. Data projects might affect public space, social interactions, personal livelihoods or even affect civil rights. To be transparent in data projects means to be able to explain the data set and its origins, the models and algorithms used to turn data into actionable information. Accountability means to take responsibility for the data collection, the analysis and models or algorithms used in it. It also means providing the necessary information which enables political parties, citizens and experts to deliberate.

Problems with transparency: Models and algorithms can become very complex and often require advanced knowledge of mathematics, statistics and data science. With transparency we do not mean to ‘translate’ this to the accessible language of a common user but to provide access for critical inquiry by experts in terms of accountability. A further problem might be that being too transparent about a dataset might give too much information to so-called criminals.

Questions

18a How transparent can you be with the public about your project?

18b Is there a danger of public outrage?

PRIVACY

Explanation

Privacy is protected by law. New regulations carry steep fees for violating privacy or leaking private information. Even if there is a general feeling that people carelessly void their right to privacy by signing up to social media services or openly sharing intimate aspects of their private lives, this does not mean that the right to privacy is losing its importance. Privacy remains essential to democracy.

Questions

19 Are sensitive data actively involved in your project?

20 Do you have insights into the private sphere of citizens?

21 Does the dataset allow insights into the personal communication of citizens?

22a Have you checked PIA (Privacy Impact Assessment?)

22b Have you had contact with a privacy officer?

BIAS

Explanation

Biases are severe issues in data analysis. A biased dataset, model or algorithm produces results that differ from the reality they aim to describe. When datasets are interpreted, certain tendencies might be pushed by those who collect data, analyze it, store it and make decision based on those data.

Example: Confirmation bias:

We all prefer to be surrounded by opinions and ideas that are similar to our own. This is the reason why many people have friends with similar viewpoints and tastes like themselves. This phenomenon was described by the behavioral psychologist B.F Skinner and is called cognitive dissonance. This mode of behaviour leads to the tendency of ignoring opinions that are not similar to our own, even though they might be very valid and important. This tendency can cause problems in the usage of data, because important outside views, different interpretations and concerns might be missing or not be heard.

Questions

23a What outcomes are you expecting personally?

23b What are other team colleagues expecting?

Example: Ingroup Biases:
Very similar to the above mentioned bias is the tendency of people to agree to the dominant opinion within the group. In case someone has a different opinion, a bad gut feeling, or a point of view different from the rest of the group, those people tend to be quiet and not communicate, because the fear of being wrong, or saying something stupid. This type of biases is highly problematical in ethical considerations concerning data because important insights of group members might be missing that could prevent negative outcomes.

24a Do you have a vague feeling about this project?

24b What do you fear? Discuss with your team members.

Example: Selection bias:
The outcomes of your data collection, visualization or interpretation might be influenced or misleading due the information you were collecting in the first place. People might be missing or are proportionally overrepresented. What looks like objective knowledge might be influenced by the type of data. Random sampling, control groups (if possible) and debates with your team might minimize the risk of selection bias.

25a Is the sample a truthful representation of the population?

25b Who is missing or invisible in your dataset?

Questions

26 **Are you gathering the right informations for your goal?**

27a **Does your decision changes thinking about long term effects? Why?**

27b **Can you imagine a future scenario in which your current decision might matter?**

INFORMED CONSENT

Explanation

Informed consent describes the approval of a person to provide information or to participate in a research project. It also means that the person has been ‘informed’ about the objectives of the research, its procedure and its implications for the participant. Informed consent has been developed as a standard for research in the wake of the cruel and illegal research experiments conducted by the Nazis in concentration camps. The principle of informed consent was implemented to avoid harm the human subject, for example in the Nürnberg Code of Conduct, which was developed shortly after World War two. The main aim of this code was to hinder future medical researcher to repeat the cruelties committed to judes. Doctors are forced to inform their patients about possible outcomes of their treatment, further research that might harm subject has to be limited. This document remained influential in the bio ethical context but also influenced current debates about informed consent. Within the medical context the relevance of informed consent is also strongly debated. It is not fully clear how to balance more kantian approaches of informed consent, by protecting individuals per se, or if it would be better to approach it more from an utilitarian perspective, by focusing more on the best for the many. (for more information go to p. 40) This debate can be for example be found in the question of how to deal with research with placebos, where it was asked if it would not be legitimate to do research without informing patients for the sake of science.

Questions

28 How do you inform people that the data is used?

29 Do citizens have a choice to opt out?

TAKING A DECISION

TAKING A DECISION

BACKGROUND INFORMATION

Different ethical perspectives with different outcomes

The question what the right thing to do is a complicated endeavor and different theories provide different outcomes.Moral theory aims to provide systematic answers to this questions about what one shall or shall not do. In the following the most important "schools"are listed. Consider that also within this different school sub debates and competing debates can be found. The following is an overall overview and can be understood as a general introduction into moral theory.

Moral relativism

"The Rightness of an action depends on the type of society you live in"

According to this view, ethical decision are social constructs. We as a community have decided what is the good and what is the bad but we could theoretically also decide the opposite. For moral relativists moral codes are entirely dependent on the moral code of a culture, which is understood as the sum of individual norms and values.An action that is considered morally wrong in one part of the world, might be considered morally right in another part of the world. Eating pork might be for example considered morally wrong for muslims whereas for christians it is not considered problematical. This view is opposed to theories that state the existence of ultimate principles that are universally right.

Connecting moral relativism to data ethics

The concept of privacy has a long western tradition. Western concepts of privacy focus on the individuals and its right to have intimacies with family and friends, secrecy and hidden correspondence. The main focus of a western concept of privacy is the individual and what belongs to this individual as a person. When we look for example at confucian traditions individual privacy (Yinsi) can be translated as "the hidden or bad thing"".

A culture however, is subject of change, and although the privacy is woven into our traditions, a cultural relativist might argue that the modern cultures have no need for such a concept and that, in fact the relevance of privacy is already declining. After all, many citizens of different nations suggest that they would not mind governmental surveillance policies because they might think that they have "nothing to hide" anyway. Further, companies and services such as facebook seem hardly to suffer from a loss of users even though they have controversial privacy policies. The disappearance of such applications of the concept of privacy might be an argument for the existence for cultural relativism.

A moral relativist might argue that in order to determine if something is a privacy infringement will depend on the basic moral code of their culture.

Critique on moral relativism

Critiques have focused on the difficulty to define what one's culture is. Can we for example say that there is something like a dutch culture? Or are we living in a midwestern culture? Or is culture more related to religious background?

Another frequently mentioned critique is that one person can also have multiple cultural backgrounds. How for example about a person that was brought up in the netherlands but has parents that come from china?

A moral relativist would respond that the difficulties of defining culture doesn't deny or even harm the claim that morality is cultural. Just because culture is complex doesn't mean it's not constitutive of our morals and values. Many philosophical arguments can be raised against moral relativism. First, if moral relativism contends that morality depends on culture, it ignores the possibility that aspects of it may be traced back to human nature. Privacy is an example of this; it may be that a desire for privacy is an evolutionary or biological aspect of human nature and therefore should not be ignored. If moral relativism allows nature to play such a role in moral theory, the position is compromised, because facts relevant to moral theory can be found in common human nature, regardless of cultures.

Relatedly, arguments have been made that there is no empirical proof for moral relativism, and that acts that we consider evil and other cultures consider good are ultimately understandable from either perspective. The classic example is that of a tribe who kill their elders when they reach the age of fifty. This seems evil to us and seems good to them. However, it seems good to them only because they believe that one retains one's body in the afterlife, thus it is after all a conception of morality related to other beliefs about the world that is understandable by us. There is therefore no hard kind of moral relativity: differences between the moralities of cultures can be explained by reference to their beliefs and histories, but not necessary by a reference to a different moral framework.

Utilitarianism

" The best for the most."

Utilitarianism is focusing mostly on the consequences of an action that bring the greatest amount of value for the most involved. Such a value can be for example the welfare or wellbeing of the most. The right thing to do is when an action as an higher utility than an other action that could have been performed instead. What exactly is understood as welfare or utility is described differently among different scholars. For Bentham and Mill, for example Welfare is identical with happiness understood as pleasure and the absence of pain.

Connecting utilitarianism to data ethics

Values like the public security and privacy frequently come in conflict with each other. Dilemmas occur if one should put more value on the one or the other. From a utilitarian perspective welfare of the many will be favored over the privacy . This argument could be also taken around if one states that privacy is necessary for wellbeing. Then a utilitarian perspective could also lead to other outcomes. Utilitarian would try to make a balancing of points in favor and points against it with focus on welfare. This is frequently called a "cost and benefit analysis". Important to know is that according to this view each individual who is involved in a certain case counts the same, that means no special treatment can be given to specific groups or people. That might cause that specific needs are overseen. To increase the overall utility and welfare of the most it might happen that minorities might be ignored. Similar arguments were made in the medical context, where it was said that a person might be sacrificed for example to be organ donors once more people could benefit from it.

Critique on Utilitarianism

A common critique of utilitarianism is that the idea of utility for the most is not really practical or useful. How can you determine what would be the best for the most one would have to know all positive and negative possible outcomes in order to balance them correctly. One possibility for utilitarianism is to introduce rules that optimize the calculus in most circumstances (so called rule utilitarianism). But this argument leads us back to the original problem, which is weighting public security against privacy as two values, which is difficult to weight and quantify. However, is such a deliberate process is done properly (e.g transparently or via legitimate political institutions), utilitarianism might serve as a guiding moral theory in developing rules that might help treating personal data ethically.

Virtue ethics

"How would a good person deal in this situation?"

Other than other theories, virtue ethics is not focusing on the question "what is the right thing to do?"but more on the question "what kind of person do I need to be to do the right thing?"'. What kind of character do I need to have to be capable of taking good decisions. This school goes back to Plato and Aristotle, who started to think about what kind of qualities of the character (virtues) are needed. Such virtues can be for example to be a honesty. To be a honest person would then imply to tell the truth and avoid telling lies. Other virtues could be courage, generosity, temperance, truthfulness, wittiness and friendliness. When those virtues are developed practical wisdom will cause that the person has developed all the skills that make her capable to take the right decision.

Connecting virtue ethics and data ethics

Professionals working with technology such as programmers and others, have frequently an solution oriented focus. This focus leads frequently to a dominance of utilitarian perspective. In the time of (big) data this might be difficult to achieve and scholars have argued that a good ethical framework for this new challanges might be virtue ethics. By that is meant that the focus is not anymore so much on the concepts of privacy and informed consent to name a few but more on the question what kind of setting and data awareness might be needed that allow professionals working with data to take responsible decision. An virtuous data analyst would therefore have for example virtues such as 'respectfulness of the sensitivity of personal data' and 'prudence and selectivity in communicating and sharing such data.' Persons can be trained into such virtues via training programs, or can be selected for these character traits.

Critique on virtue ethics

One complaint about virtue ethics tries to give some guidance for right conduct by imagining what a good person would do in this situation. This role model has developed certain virtues herself but mere imitation might not be enough for someone, who has not yet developed such virtues herself. The first objection is therefore that virtue ethics fails to give real guidance. Another struggle of virtue ethics is that it is really difficult to explain why certain character traits are virtues and others are not.Aristotle defined several virtues that are controversial, and there are cardinal virtuous (i.e. of the church as well), which begs the question: which virtuous are relevant for this particular post in mind? This is not always clear.

Another argument against virtue ethics is that it has a blind spot, or may be considered naive, towards the way in which institutions or businesses usually function, which is via organisational hierarchies. Does every employee need to be virtuous? Or do we need virtuous managers, to whom the employees need only comply? Such questions lead us away from virtue ethics itself and complicate the application of this moral theory.

Kantianism

"There are overarching principles that should guide our action."

Immanuel Kant developed a famous moral theory that is based on what he calls the categorical imperative, which is that which we must always do. Because humans are rational they can understand and derive from the categorical imperative certain rules and principles that should govern their actions. The most important formulation of the categorical imperative is that every action we commit to must be universalizable, which means that it should in principle be possible and desirable if everyone acted this way. Other formulations of the categorical imperative involve the respect for the dignity and the autonomy of human beings, meaning we should not treat others as mere means to our ends or manipulate them and take away their ability to act as they please.

Connecting to data

Data analysis promises to increase the quality of services that can be offered by a company or a government. At this level it is still relatively experimental how certain models are developed, tested and used. Following the categorical imperative could mean that even if a certain data practices sounds promising in terms of improving the public transport by tracking private mobile phone data to see movement streams of citizens in the city, it can not be done. The reason for this could be formulated as follows:Since the use of private data without the consent or knowledge of the individuals involved would violate their autonomy, it should never be done.There are ways in which these individuals might have their autonomy respected by for example allowing them to give consent or making the use of such data transparent, but currently no such policies are in place.

Critique on the kantian perspective

Some critics have argued that the kantian categorical imperative is such an abstract principle that it is not helping to guide particular situations. Furthermore, it was criticised that it might not take context relevant information that might be necessary for a particular situation. It is said that the categorical imperative is formulated in such an absolute way that it is not flexible enough to react towards situations where for example lying or stealing are not wrong things to do. Kantian moral theory might for example be too insensitive towards the particularities of certain specific practices, such as weighing privacy versus security, or how personal data collection relates to the autonomy of a person (although modern scholars have developed more nuanced views that are more readily translated to practice).

Moral particularism

"We can only decide in particular situations what the right thing to do is. "

This aspect of moral theory emphasizes that in order to evaluate if an action is morally right or wrong one has to look at facts that hold in the context. Such facts can be for example the situation of particular agents, a particular time or a particular technology. Particularists challenge the idea that a rule or principle can be found that could guide what to do. The particularist claims that the rightness and wrongness of an action is entirely dependent on the context. That means that an action is morally right when the situation and the context make a certain action necessary. Further the moral particularist does not belief in (universal) moral principles that can guide action; at best they believe moral principles only apply in comparable situations.

Connecting moral particularism to data ethics

From this perspective data practices would not be evaluated if for example informed consent is needed from the participants in general but only in this particular situation. So the question would be not so much any more : what should be the responsibility of a municipality or a company in general dealing with such issues but only case based.In practice, this means there would be much respect for the differences between cases and communication and deliberation about such cases would need to be facilitated.

Critique on moral particularism

There are two lines of criticism towards moral particularism. The first one says that without the guidance of ultimate principles nothing goes. People would then not have a reason to limit their choices and take morally decision. This critique is mostly focused on the motivation of people to act morally.

Another critique states that rationality needs to be concise. There is a danger when we only focus on each situation separately there is no overall consistency. When we for example think of someone hurting another person, we might have the difficulty to explain why this is wrong. The argument there is that without a clear notion of why hurting others is always wrong , we might fail to make a clear argument in this particular situation.

TAKING A DECISION

look through the questions you just have answered. Also check the section about possible benefits and problems of your project.Point someone in your team to be the advocat of the devel. Following concepts can help.

- Freedom of choice
- Freedom of speech
- Mutual respect
- Trust
- Diversity
- Creativity
- Peace and the good life

- **Do you treat people equally?**
  Is your approach proportional?

- **Which result would be the best for the most people involved?** *(utilitarianism)*

- **How would a good person react in this situation?** *(virtue ethics)*

- **Do you respect the autonomy of people involved?** *(kantianism)*

- **What are ethical important points specific to this situation? What makes them ethically relevant?** *(Moral particularism)*

Communication strategies

This section provides some questions that might be useful as a starting point for the development of communication strategies.

Leaks:

- What type of information was leaked?
- Was private information involved?
- Was the leak within the institution or was the leaked information publicly accessible?
- How to communicatecan technical problems be communicated within the organisation?
- How can technical problems be solved with certainty?
- Do you know how to repair the leak?

Public concerns:

- How can the public be contacted?
- What information is relevant for the public?
- How to communicate privacy issues with the public?
- How to communicate with the media? What strategies do you have for doing so?

Contact the privacy manager! Communicate with all involved team members!

Report

Clarifying the context

Answersheet

- **Project name, date, place**

- **Participants**

- **What is the project about?**

- **What kind of data do you use?**

- **Who might be affected?**

- **What are the benefits of this project?**

- **Might there be any problems with your project?**

Report

ALGORITHMS

1a Is there someone in the team who can explain how the algorithm in use works?

1b Can you communicate that with the public?

Answersheet

1a YES NO — *Go to action point.*

Who?

1b

ACTION POINT

Who will answer this question?

Report

SOURCE

2 **Where did your source get the data(sets) from?**

3 **Have you checked the quality of the data(sets)?**

4 **Is there a ‘best before’ date for this specific data set?**

Answersheet

2

3 YES — NO — *Go to action point.*

How?

4 YES — NO —

How do you delete these data

ACTION POINT

What questions are not directly answerable?

Who will answer this question?

Report

ANONYMIZATION

5 Are the data anonymized or pseudo-anonymized?

6a Have you tested the anonymization?

6b Who is in possession of the encryption key?

Answersheet

5 YES NO — *Why not?*

How are you going to do that?

6a YES NO

6b

Name:

ACTION POINT

What questions are not directly answerable?

Who will answer this question?

Report

VISUALIZATION

7 **Are the data or the produced results suitable for visualization?**

8 **How could this visualization look?**

9 **What would be a different interpretation of this visualization?**

Answersheet

7

8

9

ACTION POINT

What questions are not directly answerable?

Who will answer this question?

Report

ACCESS

10a Who has access to the dataset?

10b How is the access monitored?

Answersheet

10a

10b

ACTION POINT

Who will answer this question?

Report

OPEN ACCESS & REUSE OF THE DATASET

11a Are parts of the data suitable to be reused?
If yes, which potentials do you see with reusing data?

11b What dangers do you see with reusing data?

12 What dangers do you see with reusing data?

Answersheet

11 a YES NO — *Why not?*

which potentials do you see?

11b

12

What actions will you take to minimize risks?

ACTION POINT

What questions are not directly answerable?

Who will answer this question?

Report

RESPONSIBILITY

13 Which laws and regulations are applicable to your project?

14a Can you name a responsible person in your project?
14b Who will take care if something goes wrong?

15 Is there a danger that particular people or groups could be discriminated by your project?

16 Assess who would be suitable partners for your project.

17a In case something goes wrong, are there any commu nication strategies?
17b Who is responsible for preparing those strategies?

Answersheet

13

14a

14b YES NO — *Go to action point.*

What are the responsibilities?

15 YES NO — *Go to action point.*

Which?

16

17a YES NO — *Go to action point.*

Which?

17b

ACTION POINT

What questions are not directly answerable?

Who will answer this question?

Report

TRANSPARENCY/ ACCOUNTABILITY

18a How transparent can you be with the public about your project?

18b Is there a danger of public outrage?

Answersheet

18a

18b

Report

PRIVACY

19 Are sensitive data actively involved in your project?

20 Do you have insights into the private sphere of citizens?

21 Does the dataset allow insights into the personal communication of citizens?

22a Have you checked PIA (Privacy Impact Assessment?)

22b Have you had contact with a privacy officer?

Answersheet

19 YES NO

Which?

20

21

22a

22b

ACTION POINT

What questions are not directly answerable?

Who will answer this question?

Report

BIAS

23a What outcomes are you expecting personally?

23b What are other team colleagues expecting?

24a Do you have a vague feeling about this project?

24b What do you fear? Discuss with your team members.

25a Is the sample a truthful representation of the population?

25b Who is missing or invisible in your dataset?

26 Are you gathering the right informations for your goal?

27a Does your decision changes thinking about long term effects? Why?

27b Can you imagine a future scenario in which your current decision might matter?

Answersheet

23a

23b

24a

24b

25a

25b

26

27a

27b

Report

INFORMED CONSENT

28 How do you inform people that the data is used?

29 Do citizens have a choice to opt out?

Answersheet

28

29

Report

DELIBERATION

- **Does the project meet the standards of good governance and responsibility?**
- **Which outcome is the best for the most involved subjects, the city and its residents?** *(utilitarism)*
- **What would a person you want to be do in this situation** *(virtue ethics)*
- **Does your approach respect the autonomy of all subjects who are involved?** *(kantianism)*
- **What are problems particular in this project?** *(moral particularism)*
- **How can we contribute in a way such that the following values are respected?**

  • Freedom of choice
  • Freedom of speech
  • Mutual respect
  • Trust
  • Diversity
  • Creativity
  • Peace and the good life

Answersheet

Report

DELIBERATION

- **What are possible next steps?**
- **What did you decide, and why?**

Answersheet

DATE _____ **PLACE** _____

Colofon

Concept and development
Aline Franzke, MA

Supervision
Mirko Tobias Schäfer, Ph.D.

Commissioned by :
Gemeente Utrecht

Design:
Redmar Fransen

Illustration:
Sammy Hemerik

DEDA version 1.0,
released March 2017,
check for updates:

www.dataschool.nl/deda

Utrecht, 2017.

DEDA version 1.0. While DEDA is used we learn how to improve it. Those changes will be implemented in future versions. If you have used DEDA and you think it can be approved, please feel free to share your thought with us. Send

Acknowledgements

Developing the Data Ethics Decision Aid would not have been possible without the support, help and the feedback of others. We express our sincere gratitude to Martin Jansen who is steering the efforst for data-driven public management at the City of Utrecht. Without his commitment to responsible data practices, DEDA would not have been materialized. The input of data analysts and project managers was essential. We thank Carolina Kaitjily, Floris Fonville, Geerte van Rije, Gregor van Raalte, Hans van Impelen, Linda Scheelbeek, Ling-Po Shih for their valuable feedback, informative contributions and for participating in focus groups. Observing and inquiring the actual data practices of municipal data analysts informed and shaped DEDA. We indebted to a group of students from the Utrecht University's Applied Ethics programme: Aline Franzke, Luca van der Heide, Bas van der Kruk, Julius Bertasius. We also thank Gerwin van Schie, Iris Muis, Maranke Wieringa, and Christl de Kloe for feedback and copy editing, the entire team at Utrecht Data School, the participants of the ethics workshop at the Etmaal 2017 Conference at Tilburg University and the students of the RMA course Digital Ethics 2016-2017.

UTRECHT DATA SCHOOL

Gemeente Utrecht

Universiteit Utrecht