Ethics does not belong on the wall
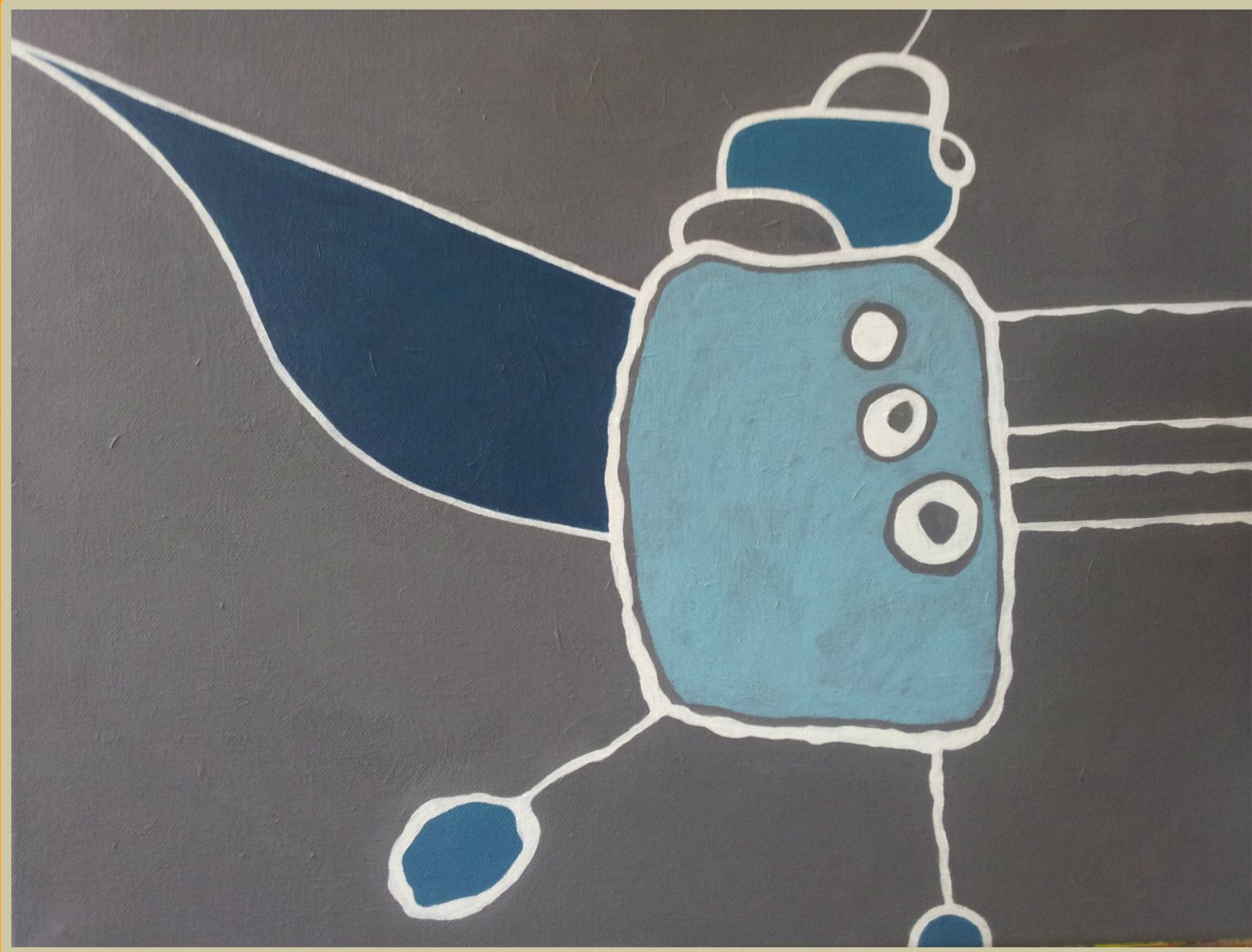

'Almost free' | acrylic on canvas| painted by Geonovum colleague Paul Padding | Studio Paul Padding | paulpadding.com

Current affairs

The corona virus is spreading around the world at lightning speed. Naturally, virologists and epidemiologists dominate the public and scientific debate. But this crisis is not only about DNA and rNA. Where people are and how they move hugely influences the spread of the virus. Therefore, information about location and travel is seen as an important tool in the fight against COVID-19.

However, the great potential of (personal) location data also begs a number of questions: What does it mean for our privacy and democratic values? Now and in the long term? Is the price not too high? It is crucial to collectively ensure that data and apps are deployed in a responsible manner.

We would like to make a valuable contribution to this. We have a common responsibility to ensure that the technology we develop and the data we use serve people, not the other way around. We must respect the individual and our shared public values in developing and rolling out applications that rely on personal location data.

This consultation version of an ethical framework has the potential to inspire data users, initiators, executive teams, clients, directors and supervisors and help them collect, use and test personal location data responsibly.

On behalf of the knowledge platform 'Wise with Location': Enjoy!
Rob van de Velde, director Geonovum

Wise with Location: https://www.geosamen.nl/wijsmetlocatie/
Locatie Pact: https://www.geosamen.nl/locatiepact/

GEONOVUM

GeoSamen
wijs met locatie

Respond via: https://www.geonovum.nl/over-geonovum/actueel/consultatie-ethische-referentie-toepassen-locatiedata

Ethics 'by design'

The kick-off of an ethical framework for the collection and use of personal location data.

Consultation version 22 April 2020
Produced by: Geonovum, Frank Verschoor, Emily Daemen

Context

The Location Pact and the knowledge platform Wise with Location were initiated as a consequence of an inspirational study trip to the US at the end of 2019. On behalf of the Ministry of Internal Affairs, the pact and platform both facilitate a congregation of experts to explore the possibilities of: a national digital twin; artificial intelligence in the geo sector; and ethics in the use of location data. All three operations are implemented and directed by Geonovum. (Read more about the study trip via https://bit.ly/3ahK9Vu)

Workshop Ethics and Location Data

Experts analysing ethics in the use of location data, examine where and how ethics and location data can offer support to those involved with the practical use of data. A key objective is to uphold and strengthen the quality, added value, impact and acceptance of data and data-related products. In order to do so, we need to understand which ethical values, -principles and -actions can be used as positive and valuable principles when designing data projects.

The Workshop Ethics and Location Data is facilitating the development of a method that ensures ethics are always part of the conversation and process in the data work environment. This method is being developed based on:

1. An inventory of existing ethical codes and questionnaires (this is available and will be sent as an attachment to this document)
2. Three workshops on use cases concerning: influencing behaviours, criminal undermining and debt counselling.
3. A series of short stories about people's individual experiences.

The Workshop Ethics and Location Data is collaborating with (but not limited to) the National Police, Land Registry, Ministry of Internal Affairs, VNG, TU Delft, RIVM, Province of South Holland and other partners from the field of location data and ethics.

Objective

The purpose of this ethical framework is to inspire data users, initiators, executive teams, clients, directors and supervisors, and to provide an additional tool for the responsible collection, use and assessment of personal location data.

Personal location data include all data that shows where people are and how they move. Traceable or not. Data from mobile apps for instance. Personal location data provide information about a person's privacy and require very careful and reserved handling. This means that it is not simply a question of following legal requirements: ethics goes beyond the minimum standards laid down in the law.

In addition, this framework should be seen as an offer for further discussion with stakeholders within the geo-professional field, the government, knowledge institutions and the business community. These stakeholders could benefit from conversations about the development and use of shared ethical standards for the responsible use of location data.

All ethical values specified and developed in this document are also relevant for non-personal and non-location-related data.

Status of this document

This is the first consultation version of an ethical framework for the use of personal location data. This document still needs to be enriched with examples, reflection and additions. Nevertheless, the value and practicality of this emerging ethical framework can already be discussed and tested in the professional field.

The next big step will be to identify how this framework can actually become a 'living' product for the benefit of the community.

Ethics 'by design'

The kick-off of an ethical framework for the collection and use of personal location data.

Consultation version 22 April 2020
Produced by: Geonovum, Frank Verschoor, Emily Daemen

Approach

This ethical framework was compiled on the basis of: an inventory of eleven ethical codes; four guidance instruments (from various organizations, companies and government bodies); and the results of the Workshop Ethics and Location Data. The workshop was about 'Influencing behaviour using personal location data'.

The analysis and coding of the inventory of ethical codes led to the ten core values that serve as a structure for the ethical framework. Each ethical code consisted of distinct rules, which we categorised into our core ethical values. In this draft framework, each core value is supported by a compilation of the most relevant rules for the use of personal location data.

The second source for the ethical framework is the aforementioned workshop. Participants in this workshop discussed and documented ethical issues and concrete solutions for each scenario, based on the core ethical values.

All issues, possible solutions, notes and other results from the workshop were also analysed, coded according to the ten core ethical values and assessed for usability for this framework.

The number of times each core value was mentioned by participants, determined the order in which the ten core values are presented and discussed in this framework.

Structure

The ten core ethical values form the main structure of this framework. Based on this structure, each core value is further developed with the insights from the workshop Ethics and Location Data and the inventory of ethical codes.

Incomplete draft

This is a consultation version. It is not yet complete, as only the first four core values have been enriched with 'reflection and additions'. Eventually, real-life examples will be included in the framework to provide a practical application of the core ethical values.

Appendix

The inventory of the eleven ethical codes and four ethical guidance instruments are sent and published separately with this framework.

Acknowledgment

Special thanks to the Johan Cruijff Arena, Geodan and DAT.Mobility for their contribution to the workshop, for the open input of their practice and experiences with the collection and processing of personal location data. Also many thanks to the USBO (Utrecht University) and The Green Land for co-designing the working methods, support and recording the results. And last but not least many thanks and appreciation to all participants in the workshop. For the valuable input and insights that have fed this framework.

Ethics 'by design'

The kick-off of an ethical framework for the collection and use of personal location data.

Personal location data

Personal location data are information about the location of a personal device, such as a smartphone. Personal location data are in the top 5 of big data, and are widely collected by platforms such as Google, Facebook, Uber and Snapchat mostly for marketing purposes. The demand for people's location data for social purposes is increasing. Examples in which this data is considered useful include traffic management, tackling crime, crowd control and 'smart city' applications for example.
In all these examples, the data are used with the aim of influencing people's behaviour.

Platforms entice users of a device to turn on the location service of an app, and then apply various techniques to track the location and use of the personal device. Thereby identifying a person and their behaviour. The data can be collected in 3 ways:

1. location data observed by technology, for example with GNSS / GPS, cell towers, Wi-Fi and Bluetooth access points;
2. technology-derived location data, for example by combining different data sets such as personal agendas, telephone lists and addresses;
3. location data given by the user, active or passive and consciously or unconsciously, by providing a location indication (geotag) in a tweet, or with a photo on Facebook or Instagram for example.

With personal location data, platforms can answer the following questions about a person:

✓ Where are you?   ✓ Where have you been?

✓ What is there?   ✓ What did you do there?

✓ What are you doing there?   ✓ Where are you going next?

✓ Who are you there with?   ✓ What will you do there?

Personal location data, continued

"Data that are completely anonymised and cannot be traced back to a person are not personal data." Aleid Wolfsen (Personal Data Authority), on Op1, April 3 2020.

Summarised based on Geonovum Reporting Exploration of location data and social platforms (available in Dutch only):
https://www.geonovum.nl/uploads/documents/2017%20Rapport%20locatiegegevens%20en%20platforms.pdf

Quick scan location determination for COVID-19

With the quick scan Location Determination for COVID-19, Geonovum has taken the first steps to create an overview of how location data could be collected from a mobile phone and to what extent that location data can then be used for various purposes. Read more and use the quick scan via (available in Dutch only):
https://www.geonovum.nl/over-geonovum/actueel/quickscan-locatiebepaling-voor-covid-19

Ethics 'by design'
"The start is the hardest part!"

Ten core ethical values

The analysis of our inventory (attached to this document) yields ten core ethical values. Which, depending on the data issue, must be implemented to data initiatives in varying degrees. Ideally these values are taken into account and used directly from the start of a project or process (Ethics by design). The inventory of ethical codes shows that everyone has their own classification and formulation of values. However, the codes are universal in content. We have reduced this content into ten core ethical principles:

1. Purpose compatibility, purpose limitation and a focus on social value
2. Active transparency
3. Data holders' and data users' accountability and responsibility
4. Engagement of relevant stakeholders
5. Protection of privacy and security
6. Establishment of and compliance with minimum ethical standards
7. Compliance with the law
8. Preservation of the quality of the data
9. Reflecting reality and inclusivity in data
10. Respecting and protecting human dignity

The above order is not particularly significant. The values are ordered based on the number of times they were mentioned and described by the participants in the workshop Ethics and Location Data. In practice, the core values ought to be applied carefully to each situation in order to accommodate for discrepancies: every circumstance is unique.

Ethics 'by design'

The core values mentioned apply to every process, project, product and service in which (location) data are used and processed. It is imperative that these values and their underlying principles, ensuing actions and agreements are added to the agenda and shaped from the start of a process, project or development of an information product or service. They must also be checked during the life and execution of (parts of) the process.

Assume that every assignment is unique and that ethics always needs to be customised. Naturally, when translating the ethical values, you benefit from 'comparable situations'. But ethics is about the small differences and nuances. Copy paste is waste when it comes to ethics. It is important that we continue to repeat this and train our ethical skills so that we can apply data even more effectively and allow solutions to truly serve their purpose.

Organise ethics in such a way that not only those who are directly involved, but also other stakeholders, know which questions to ask in order to see beyond the pretty picture. The way to achieve this is to involve them straight from the start. Especially when substantiating the ethical aspects. This reference is intended to support this process.

1. Purpose compatibility, purpose limitation and a focus on social value
"It's all about simplicity!"

Insights from the Workshop Ethics and Location Data

1. First talk to the target group, before starting with organising the data. Communicate about the goal, the method that ensures the relevant data contributes to this goal and how you would like to carry out this method. This allows the people in the target group to indicate what they find important themselves, where the boundaries are and the point at which they would like to exercise some influence. You could organise a panel that serves as a sounding board during the operational period of the project, for example.

2. It is essential to identify one clear and simple social purpose and to only organise the minimum necessary data that serves this purpose. Describe the boundaries of maximum usage of this data and include the risks that the meeting of these boundaries entail.

3. Those who organise data for more than one purpose will find themselves on a slippery slope. Especially when it concerns sensitive data. It is easy to imagine for example that new goals are easily added, the process becomes diffuse, more people join, more data are combined simply 'because' and it becomes never ending.

4. Safeguard data minimisation. Preferably at the source, rather than further along the data chain. Data minimisation means that, when (personal) data are collected and processed, only the most necessary information to reach the relevant goal is organised. Data minimisation is one of the most important principles under the GDPR. We recommend applying this principle for any form of data use.

5. If data are also used for other purposes, the risk of violating people's privacy increases because the data can be taken out of its original context. The risk becomes even greater if this leads to more linkage of data to each other. This is also known as 'function creep'. The reliability and quality of the data can then no longer be guaranteed, and thus this also leaves privacy unprotected.

6. Do not only explain in advance what you expect from the data, but also define what you absolutely do not want the use of the data to lead to.

7. Collecting data for data collection should never be an objective. This leads to both bias and searched outcomes, literally and figuratively.

8. Whenever possible, let valuable insights and other returns from the data flow back to society.

What the ethical codes say about this

Collect only data that is necessary. To the greatest extent possible, reflect and plan for the data collection within strategic planning processes. <Unicef: Ethical Considerations When Using Geospatial Technologies for Evidence Generation>

Start with clear user need and public benefit: Using data in more innovative ways has the potential to transform how public services are delivered. We must always be clear about what we are trying to achieve for users - both citizens and public servants. <UK.gov: Data ethics framework>

The Data Scientists need to understand the trade-off between gathering and collecting all potential data and focusing on just the data that is likely to be used to solve a particular problem. It is expected that Data Scientist's data gathering requests are appropriate to the problem being addressed, neither exaggerated or lacking. In any case, a Data Scientist should document the reason that a particular data set needs to be gathered. <Code of ethics>

The power and peril of data analytics is that data collected today will be useful for unpredictable purposes in the future. Give due consideration to the possibility that less data may result in both better analysis and less risk. <Accenture>

Any data use must be compatible or otherwise relevant, and not excessive in relation to the purposes for which it was obtained. The purpose of data use cannot be changed unless there is a legitimate basis. The purpose should be legitimate and as narrowly and precisely defined as practically possible. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

The purpose of data access (or collection where applicable) should be articulated no later than the time of data access (or collection where applicable) <United Nations: Data Privacy, Ethics and Protection>

Data collection and -use in public space must serve the public interest and contribute to the liveability of villages and cities. Municipalities encourage other parties to put this first. <VNG Principles of the digital society>

1. Purpose compatibility, purpose limitation and a focus on social value
"It's all about simplicity!"

Example

Reflection and additions

1. Data are never the starting point. The public data process is always derived from a concrete substantive question/need based on a clear goal. This goal can be social but can also serve personal convenience.

2. Yet too often we see in practice that a broad task, small concrete question or sometimes even randomly, as much as possible allegedly relevant data is sought, collected, processed and combined. This happens for the analysis and the hope of an unpredictable and accidental discovery. Experience shows, however, that this actually costs a lot of time (on research, data preparation and -harmonisation) and yields hardly anything. Particularly lots of noise and bias, partial and sought-after connections and beautiful insights and conclusions which do not reflect reality but confirm the assumptions.

3. Pay twice as much attention than planned, to understanding and experiencing the context, origin and narrowing of the question and underlying goal (as executive data team, for example). Only work on essential questions/goals that are demonstrably relevant in the social environment (from small to large) and where 'you' have a task/role to fulfil. Also understand, question and involve this environment (see engagement of relevant stakeholders). Avoid jumping too quickly to data and solutions. There is a big difference in approach between 'doing something with data' or 'actually solving something'

4. It is also possible to terminate the project before the appointed end date and do nothing with the data, if the expected value of the project is insufficient or the data is expected to be unsuitable for its purpose.

5. Purpose limitation is one of the most important provisions within the GDPR. Purpose limitation means that personal data may only be collected and used for a legitimate, specified and explicit pre-determined purpose (Article 5 paragraph 1 sub b GDPR).

2. Active transparancy

"Transparancy is silver, reaching people is gold!"

Insights from the Workshop Ethics and Location Data

1. Make sure that people explicitly and consciously give permission for the collection and use of their (location) data. Make it very clear for which specific purpose people make their data available. Explain why and how exactly this data and its application answer the present question. Interestingly, there is a legal basis for 'consent' in the GDPR and the Dutch Data Protection Authority offers some auxiliary tools.

2. This also applies to the use of already existing sources from public and commercial data suppliers and source holders. If the people whose data is collected have not explicitly indicated that their data may be used for the new public purpose, the existing source can only be used upon request. A general 'agreement' that leaves most people with no clue what they are agreeing to (a frequent phenomenon for most applications) is not enough.

3. If usage is temporary, clarify what happens to the data after the specified period (it should be destroyed in most cases!). Once the timeframe for data usage has expired, communicate what happened to the people involved.

4. Also indicate what it means if people do not give permission.

5. Used public data (which cannot be traced back to people) must be provided in a clear and accessible overview, in the context of transparency for all involved.

6. You must be completely transparent to the people whose data is collected and used. Which data from them is collected and used, how this data is edited, with which other parties this data is shared and for what purpose. The communication must take place in a clear and understandable way for everyone. The aim is to ensure the information actually reaches these people and other relevant stakeholders, rather than simply making it available somewhere on a website.

7. Explain the entire data process, including who is working on it, what their skills are, how the governance of the data is structured. But also what assumptions have been made, which algorithms have been used and what you do not see in the data. Stimulate engaged stakeholders to have an opinion about this.

What the ethical codes say about this

You should be transparent about the tools, data and algorithms you used to conduct your work, working in the open where possible. This allows other researchers to scrutinise your findings and citizens to understand the new types of work we are doing. <Unicef: Ethical Considerations When Using Geospatial Technologies for Evidence Generation>

Except in cases where there is a legitimate reason not to do so, at minimum, the existence, nature, anticipated period of retention and purpose of data use as well as the algorithms used for processing data should be publicly disclosed and described in a clear and non-technical language suitable for a general audience. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

transparent, understandable and accountable measures on which, when, where and for what purpose data is sourced, collected and managed should be put in place when generating data in public space. This includes both manual and automated methods, such as artificial intelligence and decision-making tools <EUROCITIES principles on citizen data>

Provenance of the data and analytical tools shapes the consequences of their use. There is no such thing as raw data—all datasets and accompanying analytic tools carry a history of human decision-making. As much as possible, that history should be auditable, including mechanisms for tracking the context of collection, methods of consent, the chain of responsibility, and assessments of quality and accuracy of the data. <Accenture: Universal principles of data ethics>

Open data is an important driver of innovation, transparency and accountability. Therefore, whenever possible, the data should be made open, unless the risks of making the data open outweigh the benefits or there are other legitimate bases not to do so. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

2. Active transparancy

"Transparancy is silver, reaching people is gold!"

Example

Reflection and additions

1. Particularly in the urgency and the pressure to deliver, transparency tends to fall off the plate or undergo a quick and dirty fix so that the product can be placed in a hidden shop window. In most cases transparency is simply forgotten or not on the radar at all, while it can actually increase the value of data and services and confidence in, for example, the government organisation involved. It is therefore imperative to make someone within the executive team responsible for transparency. Furthermore, this responsibility cannot end until the relevant information truly reaches the stakeholders involved. Transparency is silver, reaching people is gold.

2. Transparency is absolutely necessary for the public auditability of organisation involved, the engagement of the relevant stakeholders and the positive development, quality and acceptance of the product or service. The analysis of the output from the workshop and the ethical codes gives a wide range of 'to do's' in order to shape transparency. If you fully implement all the steps, there is a risk that transparency prevents you from seeing the wood for the trees. This can be prevented by involving the main stakeholders in shaping transparency. This can be the citizens whose data are used, but also administrative and official supervisors. Fulfil their minimum requirements to ensure they can perform their auditing, participating, thinking or whatever role.

3. Transparency is not only about actively showing the process, the approach, tooling the data and other facts. It is also about sharing relevant doubts, considerations, dilemmas, vulnerabilities and limitations in the data and the conclusions. Especially with critical and sensitive issues.

4. Don't turn transparency into a biased marketing or political tool. Arrange it as independently as possible, in accordance with the needs of the main stakeholders.

3. Data holders' and data users' accountability and responsibility

"Know-how and show how!"

Insights from the Workshop Ethics and Location Data

1. Ethics require a new role in the data organisation and within executive teams.
2. In order to ensure (personal) data are managed responsibly, make an independent third party a trusted supervisor who monitors compliance of ethical and other relevant agreements.
3. In addition, some situations may require a second opinion from an independent party. The assessment of suppliers and their data products, for instance.
4. Do not appoint a singular person, staff department, or committee to be responsible for ethics. This removes responsibility from others and from internal processes.
5. Have a conversation about ethics both internally and (above all) with external stakeholders.
6. Ethical codes are a guide, to be carried out by the team continuously and throughout the (data) process.
7. Arrange ethics trainings for data specialists and hold each other accountable for upholding ethical values and agreements made. Every research project could, for example, be followed up with a reflection on the fulfilment of the relevant ethical agreements – particularly when external stakeholders are involved. Ask external parties to explain how they comply with the ethical agreements. Publish and verify this explanation.
8. The commissioning board, management and department managers also each have specific roles. When they are involved with research, a project or product in which personal data is use, they must be able to pose the right ethical questions from their roles' perspective. This requires training in ethical leadership.
9. An organization can also hold itself accountable for its use of data and compliance with ethical standards in an annual report.
10. Public digital space is being privatised due to commercial parties that collect data in the public space. Why are no regulations and permits issued for this data collection?
11. Actively share public data, technology and tooling, insights and further relevant information (see active transparency). Aim for participation. Accommodate for it.

What the ethical codes say about this

The Data Scientist shall document according to a standard template each and every step along the data science value chain. This shall include the elicitation of all data sources and the usage and justification of all relevant data sources, the procedures used to combine data sources and all the steps in the data transformation pipeline. This will also include the model selection, any procedures to tune the hyper-parameters, the employed procedure to test the model and the results, and finally the strategy to industrialize the model. <Code of ethics>

The Data Scientist shall never cherry pick data or a model to back a particular statement, insight or outcome. Moreover, a data scientist shall always analyze the input data in order to assess it for any indicators of previous bias of this nature. <Code of ethics>

Consider providing training or tips on potential risks and protection strategies for individuals involved in a crowdmapping exercise. Individuals should be informed about how to conduct themselves safely in the physical environment, the ethics and risks of capturing others in any photography used in the exercise, what types of sites should not be photographed or entered (e.g. certain government buildings or locations where criminal activity takes place) as well as possible online risks. <Unicef>

A risks, harms and benefits assessment should be one of the key accountability mechanisms for every use of data, and should help determine what other governance mechanisms may be needed to monitor compliance. <United Nations>

Governments have the responsibility and have to ensure citizens can have access to and manage their data as well as influence how it is collected and used. <EUROCITIES principles on citizen data>

It is recommended that a process of due diligence be conducted to evaluate the data practices of any potential third party collaborators. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

The digital infrastructure in the common space for data collection and use is readily available and accessible to everyone. <VNG>

3. Data holders' and data users' accountability and responsibility
"Know-how and show how!"

Example

Reflection and additions

1. Work with people and parties who clearly want and are able to act intrinsically ethically. Preferably from a shared ethical frame of reference. Only work with suppliers who evidently apply similar minimum ethical standards.

2. Each link in the data chain is (partly) responsible for: what is collected, processed and developed in the same data chain; for the positive and negative effects thereof; and for compliance with agreed ethical standards.

3. Directors, clients, project leaders, data specialists, suppliers, customers, supervisors and everyone else in the data chain must be able to raise ethics related questions about the data process as part of their role (structured but when in doubt... speak up!). They should be able to act and take responsibility.

4. The introduction of ethical standards has consequences for the data strategy, for the data policy, for the structure of the data organisation, for the work process, for the competences and for the roles and responsibilities.

5. In order to prevent data and technology from being used too quickly in a given assignment, data specialists (in particular) can be very critical towards their clients regarding the why, the origin of the question, the purpose, the effect on biased steering, and the assumptions that are done et cetera. And vice versa, it is the responsibility of clients to involve data specialists at an early stage in the assignment.

6. It is also the task of the data specialist to illustrate the complexity, possibilities and limitations of data to less proficient data stakeholders. This would include them in the data process and stimulate them to ask the right questions about it.

4. Engagement of relevant stakeholders
"Don't collect data about, but with people!"

Insights from the Workshop Ethics and Location Data

1. The most vital stakeholders are the people who share their personal data and are influenced in any way by the use of the data. Ethics serve them.
2. It is important to already involve citizens and other primary stakeholders in the design phase of the project. And also to give them influence as 'co-owner'. Not only to analyse and limit risks, but also to increase the quality and usability of the data and its 'application'.
3. One way to engage and empower citizens in a practical way is to conduct an experiment with them, also as part of the collective design process.
4. Do not consider the citizens solely as data generators, but also as co-creators and 'end users'.
5. Involve the Dutch Data Protection Authority (abbreviated in Dutch as: the AP) and its considerations regarding the use of personal data at an early stage in the process.
6. Particularly when influencing behaviour, the aim is to change the way people move. People will only do so, if they feel connected to the goal and the organiser. And if they receive something in return for their contribution.
7. Municipalities are an important stakeholder and vehicle for reaching and engaging citizens when tackling societal challenges.
8. The quality of ethics depends on the quality of the conversation with the right people and the critical questions being asked.
9. Collaborate with and learn from (social) parties that pursue the same goal.

What the ethical codes say about this

Decisions concerning the use of sensitive data should involve consultation with groups concerned (or their representative) where possible to mitigate any associated risks. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

Collecting and combining data may result in unforeseen insights on society or individuals. Parties collecting data in public spaces should ensure they regularly engage citizens to investigate, discuss and agree requirements for any ethical consequences of data collection and adjust their practices to prevent all forms of discrimination based, for example, on gender, age, socio-economic status, ideology, race or religious beliefs. <EUROCITIES principles on citizen data>

Governments have the responsibility and have to ensure citizens can have access to and manage their data (e.g. MyData), as well as influence how it is collected and used. <EUROCITIES principles on citizen data>

Citizens' engagement should be encouraged and organized in such a way that it persists after the launch of a project or the development of a programme. <digital society (VSNU) data principles>

Those citizens should be involved who are directly involved and affected by the relevant aspects and activities. <digital society (VSNU) data principles>

Collaborate with all relevant stakeholders to populate a risk assessment framework. <Unicef: Ethical Considerations When Using Geospatial Technologies for Evidence Generation>

Residents and users have a say in the design of our digital city. The government, social organisations and companies facilitate this. They monitor the development and the social consequences. <tada.city>

4. Engagement of relevant stakeholders
"Don't collect data about, but with people!"

Example

Reflection and additions

1. The goal is never to use data and technology. Therefore, data and solutions are never the starting point for talking to and involving the most important stakeholders: the people who need to 'move' in order to achieve a certain goal. For example: football supporters who need guidance around a stadium to avoid congestion and conflicts (crowd management); people who have to stay at home and who have to keep their distance to manage COVID-19; people who want to travel from A to B through the fastest, greenest or cleanest route.

2. If you have already made choices or come up with 'good ideas' about the use of data and solutions before involving the main stakeholders and experiencing the situation in question, then these choices and ideas are by definition biased. They steer in a certain direction and are full of assumptions that the inventors hope to see confirmed.

3. The starting point is the context, dynamics and perspective of the people who ultimately have to or want to 'move' so that a specific shared or individual goal can be reached. By actively consulting, engaging and understanding these people at the onset, before any plans are formulated or made, 'should' transforms into 'want'. Reflect their voice in the purpose specification (this tends to bring a welcoming, insightful clarification), the planning, the solution and the verification of compliance with ethical standards and other agreements. This also means that you must be transparent about how their voices are reflected and can demonstrate their involvement.

4. The nature and extent of the involvement of relevant stakeholders differs per situation and is influenced, among other things, by the impact (corporate, personal, small-business), the duration (one-off, temporary or structural), the risks (privacy, security, equality), the geographical scope (international, national, local), the domain (public, private) of the problem, the purpose and the task.

5. Protection of privacy and security

"I spy with my little eye…"

GEONOVUM

GeoSamen wijs met locatie

Insights from the Workshop Ethics and Location Data

1. Privacy is one of our fundamental rights, expressed in article 10 of the Dutch Constitution. The General Data Protection Regulation (GDPR) is the most important instrument for the processing of personal data in the EU. The Dutch Data Protection Authority is our national data protection supervisor and supports the correct implementation of the GDPR.

2. All core ethical values contribute to the protection of the privacy and security of people who share their personal data for processing and of those who are affected in any way by the use of data.

3. Give people control over their own data.

4. Make it clear with whom (personal) data is shared. Make the data chain and all operations completely explicit and transparent.

5. Give people the opportunity to check whether agreements concerning the use of their data are complied with and how.

6. Make sure that people explicitly and consciously give permission for the collection, use and sharing of their (location) data. Make it very clear for which specific purpose people make their data available. Explain why and how exactly their data and its application answer the underlying question.

7. This also applies when already existing sources from public and commercial data suppliers and source holders are relied on. If the people whose data is collected have not explicitly indicated that their data may be used for the new purpose, the source can only be used upon request. A general 'agreement', without people even knowing what they agree to, as is requested with most applications, is not enough.

8. If usage is temporary, specify what happens to the data once the timeframe of usage has ended (the data should be destroyed in most cases!). Communicate what has been done after the period of use, to the people involved.

9. Also explore 'no go areas' for data to protect vulnerable locations and communities.

What the ethical codes say about this

Data subjects hold a range of expectations about the privacy and security of their data and those expectations are often context-dependent. Designers and data professionals should give due consideration to those expectations and align safeguards and expectations as much as possible. <Accenture: Universal principles of data ethics>

Consider the privacy policies of third-party geospatial data providers (such as social media services). When using third-party data, consideration should be given to the privacy policies of the organization and their implications including (where relevant): Anonymization and aggregation of data provided by the third party; Safe transmission mechanisms for data (e.g. encryption used at all times when data is being sent from one party to the other); Whether there are clear conditions evidencing respect for individuals' rights relating to their data. This could include consent arrangements for non-operational use of data, notification of potential sharing of data (including information about with whom it may be shared), right to removal of personal data from data sets, etc. When deciding (a) whether to use the third-party data and (b) whether it is feasible or appropriate to create an MoU to ensure privacy and security in the transfer and receipt of data or analysis. <Unicef>

If deciding to proceed with a partnership to accept geospatial data from third parties, then measures should clearly be taken to publicly acknowledge the nature of any partnership and the safety measures taken to protect the privacy of those whose data has been used. <Unicef: Ethical Considerations When Using Geospatial Technologies for Evidence Generation>

Encrypt personal and sensitive data when transferred to or from any network-connected server. No de-identified data should knowingly and purposely be re-identified, unless there is a legitimate, lawful and fair basis. To minimize the possibility of re-identification, it is recommended that de-identified data not be analyzed or otherwise used by the same individuals who originally de-identified the data. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

6. Establishment of and compliance with minimum ethical standards
"Ethics does not belong on the wall!"

Strongly corresponds with core value 3. Data holders' and data users' accountability and responsiblity.

Insights from the Workshop Ethics and Location Data

1. Ethics require a new role in the data organisation and within executive teams.
2. In order to ensure (personal) data are managed responsibly, make an independent third party a trusted supervisor who monitors compliance of ethical and other relevant agreements.
3. In addition, some situations may require a second opinion from an independent party. The assessment of suppliers and their data products, for instance.
4. Do not appoint a singular person, staff department, or committee to be responsible for ethics. This removes responsibility from others and from internal processes.
5. Have a conversation about ethics both internally and (above all) with external stakeholders.
6. Ethical codes are a guide, to be carried out by the team continuously and throughout the (data) process.
7. Arrange ethics trainings for data specialists and hold each other accountable for upholding ethical values and agreements made. Every research project could, for example, be followed up with a reflection on the fulfilment of the relevant ethical agreements – particularly when external stakeholders are involved. Ask external parties to explain how they comply with the ethical agreements. Publish and verify this explanation.
8. The commissioning board, management and department managers also each have specific roles. When they are involved with research, a project or product in which personal data is use, they must be able to pose the right ethical questions from their roles' perspective. This requires training in ethical leadership.
9. An organization can also hold itself accountable for its use of data and compliance with ethical standards in an annual report.
10. Find fun ways to internally discuss ethics.
11. Work 'together' towards generally accepted minimum ethical standards. Make as much use as possible of what is already there. Experiment and learn from applying the different core values and underlying aspects to starting and ongoing projects.

What the ethical codes say about this

Aspire to design practices that incorporate transparency, configurability, accountability, and auditability. Not all ethical dilemmas have design solutions, but being aware of design practices can break down many of the practical barriers that stand in the way of shared, robust ethical standards. Data ethics is an engineering challenge worthy of the best minds in the field. <Accenture: Universal principles of data ethics>

We think upholding our Code, values and ethos in large affairs begins with ethical conduct in small affairs — in doing the right thing, every time, even when no one is looking. We also think it resides in creating and protecting a culture where we remain tough on ideas but fair with people. <Planet.com Code of ethics>

We seek to work with customers and partners who adopt similar data ethics principles. <Planet.com Code of ethics>

Governance practices should be robust, known to all team members and reviewed regularly. Data ethics poses organizational challenges that cannot be resolved by familiar compliance regimes alone. Because the regulatory, social, and engineering terrains are so unsettled, organizations engaged in data analytics require collaborative, routine and transparent practices for ethical governance. <Accenture: Universal principles of data ethics>

If third party data or consultancy expertise is to be used, ensure that a non-disclosure agreement is included in the procurement process and that experts are briefed on any relevant, organizational ethical procedures and requirements pertaining to privacy. <Unicef: Ethical Considerations When Using Geospatial Technologies for Evidence Generation>

Organizations should prioritize establishing consistent, efficient, and actionable ethics review practices for new products, services, and research programs. Internal peer-review practices can mitigate risk, and an external review board can contribute significantly to public trust. <Accenture: Universal principles of data ethics>

7. Compliance with the law
"Doubt is the best counsellor!"

GEONOVUM

GeoSamen
wijs met locatie

Insights from the Workshop Ethics and Location Data

1. Privacy is one of our fundamental rights, expressed in article 10 of the Dutch Constitution. The General Data Protection Regulation (GDPR) is the most important instrument for the processing of personal data in the EU. The Dutch Data Protection Authority is our national data protection supervisor and supports the correct implementation of the GDPR.
2. In addition, there are all kinds of Dutch laws that regulate the use of data, such as the law on open governance (in Dutch: Wet Openbaarheid van Bestuur), the law governing our physical 'living environment' (in Dutch: De omgevingswet), the law on basic registrations, the archive act, the police data act, etc.
3. An inventory and analysis of relevant legislation prior to the use of data is necessary.

What the ethical codes say about this

Open data, transparency and accountability: Appropriate governance and accountability mechanisms should be established to monitor compliance with relevant law, including privacy laws and the highest standards of confidentiality, moral and ethical conduct with regard to data use. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

Third party collaborators engaging in data use should act in compliance with relevant laws, including privacy laws as well as the highest standards of confidentiality and moral and ethical conduct. Their actions should be consistent with the United Nations' global mandate as well as UN regulations, rules and policies. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

The Data Scientist will always act in accordance with the law, developing a full knowledge of, and ensuring compliance with, all relevant regulatory regimes. Employers should take steps to raise their data scientists' awareness and knowledge of such issues. <Code of ethics>

Be aware of relevant legislation and codes of practice: You must have an understanding of the relevant laws and codes of practice that relate to the use of data. When in doubt, you must consult relevant experts.<UK GOV Data ethics framework>

Always follow the law, but understand that the law is often a minimum bar. As digital transformations have become a standard evolutionary path for businesses, governments and laws have largely failed to keep up with the pace of digital innovation and existing regulations are often mis-calibrated to present risks. In this context, compliance means complacency. To excel in data ethics, leaders must define their own compliance frameworks that outperform legislated requirements. <Accenture: Universal principles of data ethics>

8. Preservation of the quality of the data
"Of people, for people and done by people!

Insights from the Workshop Ethics and Location Data

1. If data are also used for other purposes, the risk of violating people's privacy increases because the data can be taken out of its original context. The risk becomes even greater if this leads to more linkage of data to each other. This is also known as 'function creep'. The reliability and quality of the data can then no longer be guaranteed, and thus this also leaves privacy unprotected.

2. Aside from the technical quality of individual data, the quality of data is mainly determined by the extent to which it actually meets the demand and the specific purpose. Can the data deliver what is really needed? Don't compromise on that. Otherwise, the data will drive the project and data bias lurks.

3. People are only willing to invest in prioritising high value data and safety if it has a very high use value and its purpose is considered useful. In all other cases, it is better to terminate the mission, because the risk of failure is too great.

What the ethical codes say about this

The quality of the data should be preserved. Those who use and share data have the responsibility to ensure the integrity, authenticity, consistency and accuracy of data. <EUROCITIES principles on citizen data>

All data-related activities should be designed, carried out, reported and documented with an adequate level of quality and transparency. More specifically, to the extent reasonably possible, data should be validated for accuracy, relevancy, sufficiency, integrity, completeness, usability, validity and coherence, and be kept up to date. <Unicef: Ethical Considerations When Using Geospatial Technologies for Evidence Generation>

We strive to deliver data and analytics that are accurate and documented to the best of our ability, and to ensure their veracity. <Planet.com Code of ethics>

Automatic processing of data, including the use of algorithms, without human intervention and domain expertise should be avoided when data is analysed for decision-making that is likely to have any impact on an individual(s) or group(s) of individuals to avoid potential harms resulting from low quality of data. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

Data quality should be carefully considered in light of the risks that the use of low quality data for decision-making can create for an individual(s) and group(s) of individuals. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

A periodic assessment of data quality is recommended during the data life cycle. Furthermore, it is important to establish an internal system of constant data updating and deletion of obsolete data, where appropriate and practically possible. <United Nations: Data Privacy, Ethics and Protection. A Guidance Note on Big Data for Achievement of the 2030 Agenda>

9. Reflecting reality and inclusivity in data

"Everything is different outside!"

Insights from the Workshop Ethics and Location Data

1. Even if you are not personally featured in the data, conclusions will still be drawn about you. Also if you have not given permission for your data to be collected and used.
2. People often talk about 'social data' when the personal data of residents are used. The residents whose data are used are then 'society'? Again conclusions are also drawn about those who do not appear in the data.
3. Often conclusions from a very small population are said to represent a very large population. Even if the use of the data and calculations are correctly performed, legally permitted and scientifically substantiated, the risk of unfair conclusions is still very great.
4. To what extent can we speak of social data if agreements about data collection and use are laid down in an agreement between two organisations?
5. What is the value of your permission? For yourself and for society? Do you consider influence a fair reciprocation?
6. Models never tell the truth. That is why it is very important to clearly explain and live through algorithms and models. What goes in and what comes out?

What the ethical codes say about this

The Data Scientist is responsible for questioning the data before creating any model and understanding the reasons why a particular data set have passed certain filtering criteria without overlooking those data items that didn't<Code of Ethics>

Data can be a tool of inclusion and exclusion. While everyone deserves the social and economic benefits of data, not everyone is equally impacted by the processes of data collection, correlation, and prediction. Data professionals should strive to mitigate the disparate impacts of their products and listen to the concerns of affected communities <Accenture: Universal principles of data ethics>

The quality of the data should be preserved. Those who use and share data have the responsibility to ensure the integrity, authenticity, consistency and accuracy of data. <EUROCITIES principles on citizen data>

Data quality must be assessed for biases to avoid any adverse effects, where practically possible, including giving rise to unlawful and arbitrary discrimination. <United Nations>

Consider the possibility of discrimination against disadvantaged groups that are collectively associated with particular geographical areas. Correlations between particular populations in light of factors such as their geography and the relationship between location, poverty, gender and race may result in geographical trends and predictive models that discriminate against certain persons in relation to their access to services and their opportunities. Where discrimination is a possibility, the use of geospatial technologies and data should be reconsidered and/or any resulting decision making should be carefully triangulated with other data sources. <Unicef>

The Data Scientist has a duty not to break gender, race, ethnicity, martial status, religion, belief, disability, or age equality legislation. In particular, such attributes should not place individuals at any disadvantage within models or any automated decisions. <Code of ethics>

10. Respecting and protecting human dignity
"Ethics go beyond law and science!"

GEONOVUM

GeoSamen
wijs met locatie

Insights from the Workshop Ethics and Location Data

1. Ask yourself, do we have to solve this? If so, is this the best way? Not everything has to be solved.
2. Ethics is also the translation of our democratic principles into the development and end (data) products we make.
3. It is precisely in a democracy that you want to have a transparent conversation about the utility, necessity and consequences of collecting and using data in light of fundamental values such as freedom and equality. What do we want and what do we not want?
4. Everyone involved has a moral responsibility to ensure that people have the possibility of making conscious decisions for themselves in this regard.
5. "You are no longer free in public space." Public space is being digitised and privatised, and has been placed under major surveillance by private parties. We need to think about this together and organise and regulate it properly.

What the ethical codes say about this

Human dimension: Data and algorithms are not the last word. Humanity always comes first. We leave room for unpredictability. People have the right to be forgotten digitally. This always leaves room for a new, clean slate. <Tada.city>

I commit to always put humans before business, and to stand up against pressure to do otherwise, even at my own risk. <Tech Pledge>

The highest priority is to respect the persons behind the data. When insights derived from data could impact the human condition, the potential harm to individuals and communities should be the paramount consideration. Big data can produce compelling insights about populations, but those same insights can be used to unfairly limit an individual's possibilities. <Accenture: Universal principles of data ethics>

I commit to fight for democracy and human rights, and to improve the institutions that protect them. <Tech Pledge>

We encourage the use of our data to promote inclusive human well-being and development, and work to minimize its potential for harm, abuse, violations of human rights, aggression or violence. <Planet.com Code of ethics>

Algorithms and models. Prejudices and attempts against fundamental rights: Data scientists shall not create inferred evidence that violates fundamental principles, such as presumption of innocence, etc. <code of ethics>

I commit to work towards a more equal, inclusive and sustainable future for us all, following the United Nations global goals. <Tech pledge>

Data access, analysis or other use must be consistent with the United Nations Charter and in furtherance of the Sustainable Development Goals. <United Nations>

10. Respecting and protecting human dignity
"Ethics go beyond law and science!"

Example

Reflection and additions

1. Article 1 of the EU Charter of Fundamental Rights with protects human dignity: 'Human dignity is inviolable. It must be respected and protected.' This article is also included in the Dutch Constitution.
2. A clear definition of human dignity is not given and the term is used in different ways.
3. Since 'protecting human dignity' is difficult to translate into rules and practical tools for using data, we have been hesitant to include it in this ethical framework.
4. However, the discussions surrounding the use of data and apps to control Corona confirm that "protecting human dignity" is, if possible, the most important ethical value when using personal location data. Because the potentially far-reaching but still unpredictable consequences of this can undermine the (democratic) fundamental values of Dutch society, such as freedom and equality (also in the long term).
5. Especially in situations where society is under pressure, such as during the Corona crisis, we must collectively embrace and monitor these values even more.
6. But that doesn't mean you can 'squander' human dignity in less critical situations. Also on a micro level, freedom, equality and humanity must be protected.
7. The core value of respecting and protecting human dignity functions as a safety net, so that the question 'Do we want this as a society?' never goes unanswered.
8. Because a certain development can adhere nicely to the core ethical values 1-9, but still affect our fundamental values.
9. When collecting and using personal location data, the technology used is also ethically essential. Technology often tends to be applied in an autonomous, pervasive way. To protect human dignity, technology must therefore always be viewed and used as a person's tool.