Deliverable

D1.6 Ethical Principles for using Data-Driven Decision in the Cloud (It. 2)

Project Acronym:	DUET	
Project title:	Digital Urban European Twins	
Grant Agreement No.	870697	
Website:	www.digitalurbantwins.eu	
Version:	1.0	
Date:	31 August 2022	
Responsible Partner:	GSL	
Contributing Partners:	AIV, IMEC	
Reviewers:	Susie McAleer (21C); Lieven Raes (AIV); Andrew Stott; Yannis Charalabidis, Geert Mareels	
Dissemination Level:	Public	X
	Confidential – only consortium members and European Commission	

This project has received financial support from the European Union's Horizon 2020 Programme under grant agreement no. 870697

Revision History

Revision	Date	Author	Organization	Description
0.1	7.7.2022	Tomas Pavelka, Annabel Pemberton, Marco Lauro, Simona Uhrinová	GSL	Initial structure and first draft
0.2	18.08.2022	Tomas Pavelka, Annabel Pemberton, Marco Lauro, Simona Uhrinová	GSL	First draft
0.3	23.8.2022	Tomas Pavelka, Annabel Pemberton	GSL	Second draft
0.4	30.8.2022	Andrew Stott, Geert Mareels, Jiří Bouchal	GSL, AIV, external experts	Review
1.0	31.8.2022	Tomas Pavelka, Annabel Pemberton	GSL	Final version

Table of Contents

Executive Summary

This deliverable seeks to provide an updated and final version of the "Ethical Principles for using Data-Driven Decision in the Cloud" (the 'Ethical Code of Conduct') tailored to assist in any data-based decision-making process: (i) discuss the building blocks of the ethical discourse around data-based decision making, and (ii) suggest an ethical code of conduct (ethical principles) for cities in such a context.

The deliverable also assesses and extends the previous guide provided by D1.5 and updates the background of applicable laws.

This deliverable **should not be regarded as legal advice.** Organisations' legal departments or external attorneys qualified in the concerned jurisdictions should be consulted with respect to any particular legal matter.

This document comprises of feedback on the early Ethical Code of Conduct section:

- **Summary of the general use of the early Ethical Code of Conduct** by the pilot cities Athens (DAEM), Region Flanders (Digital Flanders, formerly AIV) and Pilsen (SITMP). Feedback was collected through a questionnaire and online interviews. The questions posed to the pilots can be found in **Annex 1.**

- **Granular feedback on the usability of the early Ethical Code of Conduct sections** including personal/non-personal data, selection of data bases, risks of certain data types, documenting and communicating the use.

- **Areas of improvement that were identified** relating to the content, structure and user experience for the future presentation of the content of the Ethical Code of Conduct. In particular, while the content was up to date, relevant three main improvement areas were highlighted including additional user types in the Ethical Code of Conduct, **continuous monitoring of new epics and additions to the Preliminary considerations on ethics and the Cloud as elaborated upon in Section 3.**

Which is followed by a final version of the Ethical Code of Conduct:

- **The final version** provides an updated charter of 12 core ethical principles for making data-driven decisions in **Annex 2.** Following an introduction (including a necessary legal disclaimer), the ethical principles section elaborates on the following areas outlined in D1.5 with the addition of:

    - **Additional user roles** including investor, entrepreneur and smart city providers (inclusion of private sector).

    - **Further sections in Section 2,** including a definition of private/public data sources with reference to external publications including ENISA.

    - **Links to previous deliverables within the guidance** for ease of navigation.

1. Introduction

This deliverable is the second iteration in creating Ethical Principles (The '**Ethical Code of Conduct**') to ensure appropriate considerations are made to ensure compliance in data-driven decision making.

Alongside the final iteration of the Cities Guide (D1.4), this deliverable aims to further provide guidance for DUET in legal and ethical matters. D1.5 Ethical Principles for using Data-Driven Decision in the Cloud (It.1) began the ethical guidance, through highlighting questions which cities should consider when choosing data for their models, and safeguards to mitigate potential harm to users and a wider stakeholder group. In particular, D1.5 aimed to (i) discuss the building blocks of the ethical discourse around data-based decision making, and (ii) suggest an ethical code of conduct (ethical principles) for cities in such a context.

Similar to D1.4, the present deliverable is focused on the feedback from the DUET pilot city partners on D1.5. Based on that feedback, this final deliverable provides evaluation of the Ethical Code of Conduct, suggests improvements to be included in the final re-stated guidance and publishes the final version of the Ethical Principles.

This iteration of the Ethical Code of Conduct deliverable therefore focuses on:

- Analysis of feedback from DUET partners regarding format, content, usefulness and impact of the guidance in It. 1;
- Identifying any gaps in the existing guidance published in D1.5, assessed from the content perspective, as well as format and user experience;
- Publishing a final guidance version of the Ethical Principles;
- Further work that could be carried out.

Accordingly, Chapter 1 summarises feedback collected and assessed from Pilot city partners and DUET expert team via written Q&A and interviews over video conference.

Chapter 2 looks at the identified or perceived gaps and items for improvement/further research.

Conclusion provides an overall evaluation plus summary of the steps to be taken to increase and spread awareness of the DUET project's achievements in the legal and ethical fields

Annex 1 sets out the questionnaires submitted to Pilot partners for purposes of this deliverable.

Annex 2 contains the final version of the "Ethical Code of Conduct." This version integrates the general guidance provided in D1.5. In addition, it supplements / plugs some of the gaps identified throughout the review, even though the adjustments had to be only minor. From a legal point of view, the guide can be considered as up-to-date.

This document and the final guideline version is complementary to, and does not replace the general guidance in further detail described in deliverable D1.4 (Cities Guide to Legal Compliance for Data-Driven Decision Making).

* * *

Legal notice: Even though it is an ambition of this document to provide a useful guidance to any interested smart cities and other stakeholders out there, it is important to note that this document or deliverables D1.1, D1.2-D1.3-D1.4 (cities guide) and D1.5 (ethical principles) do not, and are not intended to, constitute legal advice to DUET partner organisations or any third parties. Instead, all information, content, and materials in these documents are for informational purposes only within the scope and objectives defined for the respective DUET project deliverables. Given that these documents got finalised at a certain point in time, information in these documents may not constitute the most up-to-date legal or other information at the cut off date. Readers of these documents and their organisations should contact their in-house team members (including their **Data Protection Officers** (DPOs)) or an attorney qualified in the concerned jurisdictions to obtain advice with respect to any particular legal matter.

2. Feedback from DUET partners

For purposes of this deliverable, we created questionnaires on the use of deliverable D1.5 within pilot organisations for the DUET project purposes. The questionnaire is reproduced as **Annex 1**.

Based on an analysis of the responses obtained, we conducted a series of interviews with each pilot city partner organisation. The responses / feedback are summarised in this chapter. As previously for purposes of drafting deliverable D1.2 and D1.4, questionnaires were fashioned as relatively open-ended, hence the responses differed in focus and granularity of information.

2.1 General use of the Draft Ethical Principles

Athens (DAEM)

DAEM is a private organisation supporting the City of Athens services. DAEM considered D1.5 as a useful document meeting the goal of providing an ethical code of conduct for cities in data-driven decision making and appreciated that the document (i) defines the roles of different actors and their rights on data processing and (ii) covers the ethics related to new technologies in the cities' decision making.

In terms of utilising Section 2 of D1.5, concerning preliminary considerations on ethics and the Cloud, DAEM mentioned that their legal department had an overview of the content involved.

As regards the format of D1.5, DAEM suggested the Ethical Code of Conduct to be transformed into a practical guided (online) questionnaire for cities playing a role of a first step "ethics-check" regarding the cities' role and technologies.

While only informational, DAEM considered the steps in Section 3 of the Guide (the DUET Ethical Principles) sufficient and specifically found the steps useful for defining the roles in the DUET landing pages and logins. Moreover, DAEM mentioned that they gained a better understanding of the differences between legal and ethical standards due to the information included

Flanders (Digital Flanders, formerly AIV)

Digital Flanders (**DV**) is part of the Flemish Government, in Belgium. DV (formerly named AIV) is a public body tasked with support in the areas of digitization of data in e-government, GIS and public information.

According to the pilot partner, the Ethical Code of Conduct met its goal of being an ethical code of conduct for the cities and its expectations relating to what they needed in Flanders during the Digital Twin implementation. DV further highlighted that the D1.5 contains a good overview of the elements necessary to take into account and, even though the Flanders open data principles (note: meaning the Smart Flanders Open Data Charter) already provide a good code of conduct according to DV, they stressed that the ethical dimension is far more at the centre in the D1.5. According to DV, Section 2 of the D1.5 is quite complete and overarches the relevant topics, including information on use of HPC and future AI.

However, DV considered (in line with other partners) that given the lack of personal data used for DUET purposes, the guidance was of less direct relevance for the Flanders use cases.

Based on the DV's experience, the traffic model-based simulations appear the most challenging from the privacy or ethical perspective, although the link to using personal data is almost absent and was already solved during the model creation. Another ethical, more general, element is in the DV's view related to the value of the outcome of combining two simulation models and the level of trust and reliability.

DV pointed out that they consulted Section 2 of D1.5 due to utilising the Microsoft Azure hosting of DUET which is operated on EU-based servers.

In terms of Section 3 DV declared that the Guide helped them gain a better understanding of the transparent and fair use of AI and computer models due to the information included in Principle 7 of Section 3 (**'Transparent and fair use of AI and computer models. Fighting the "opacity" problem.')**.

Furthermore DV stressed that Section 3.10 (**Privacy by Design**) was relevant in the case of the schoolstreet Telraam pilot. However, this pilot originated from the PoliVisu project and was only touched upon in DUET and not fully integrated into the Digital Twins.

Pilsen (SITMP)

Správa Informačních Technologií města Plzně (SITMP), as a part of the city of Plzeň, is a public company responsible for ICT of the city Plzeň (Pilsen).

Like Athens and Flanders, Pilsen also considers D1.5 meets the goal of addressing ethical issues that smart cities should consider when making data-driven decisions. This includes the most important legal norms included in Section 2. Similar to other partner pilots, SITMP has not consulted Section 2 in practice due to a lack of circumstances arising to require it. Nonetheless, similar to Flanders, Pilsen considers Sections 2 exhaustive in terms of its content and even over-inclusive relating to AI due to absence of AI-related use cases. SITMP also appreciated an overview of likely ethical considerations relating to various user types in Section 3.

SITMP again pointed out to the fact their current use cases involve only traffic data and models which are not likely to cause privacy or ethical issues. Therefore, they do not consider any of them challenging from the perspective of privacy or ethical considerations. For the same reason, SITMP has not used the Section 3 framework yet but found it helpful for future use. SITMP particularly appreciated the list of ethical principles (pages 34-36) which they considered a good checklist for future use and possible ethical issues listed at each of the user types on p. 30 – 32 being a part of Section 3.

When asked about their opinion on the ethical framework being up to date with developing DUET ambitions, SITMP especially pointed out to problematic data sharing within city authorities, citing the following answer included in Section 3: "I was not sure whether my City has data available that would help me to complete a particular task. I was unsure who to turn to to find out." According to SITMP, this is a problem of large authorities and cities that manage huge amounts of data, information, and applications in different places (websites, applications, registers, etc.), so it is sometimes difficult for employees to find a suitable data source to solve a problem, or they do not even know that such an application/site exists.

2.2.Preliminary considerations on Ethics and the Cloud

As shown in the questionnaire in Annex 1, pilot partners were asked questions based on each section of D1.5, including Section 2, the **Preliminary considerations on ethics and the Cloud**.

In response, the pilot partner DAEM suggested **providing an additional reference to ENISA**[1]. The pilot partner also suggested **drafting an additional data-sources section defining private/public data sources**, to assist with issues including checking any usage restrictions on private sources if a pilot is buying the source.

The pilot partner DV commented that Section 2 could also include **steps to know where the data will be stored physically and information about the regulation applied to the service provider** (US, Chinese law, etc.). The reasoning is practical, and that the location is becoming more contentious and necessary to disclose to be in line with data protection regulations. While exploring other jurisdictions is outside of the scope of this deliverable, appropriate steps have been added to the final iteration to provide a holistic Guide.

2.3 Improvements and changes to the Draft Ethical Principles

Pilot partners were also asked questions based on Section 2, the **DUET ethical principles - first version.**

In response Athens suggested the addition of 'investor'/entrepreneur, startup roles, while Flanders further commented that Smart city service providers (private sector) roles to be added. Both Athens and Flanders also suggested that the Ethical Principles in this section be transformed into a practical online guide / questionnaire for the cities, possibly with an inclusion on the DUET website with the help of which the cities could do the first viability or ethical check. In particular, Athens specified that each question would direct the user to another section according to their answer, e.g. "What role do you have? > the relevant section > will you collect personal data from the citizen?". Flanders suggested directing the user depending on the answers to the next relevant topics and that the outcome be a list of dedicated principles the city needs to check, leading to "likely ethical consideration" as mentioned in Section 3. of D1.5

In relation to content, Athens also suggested that guidelines on legal and ethical procedures when integrating DUET with other existing systems or newly developed systems for its further development could also be included. Pilsen mentioned that the most important (additional) content would be the monitoring of new epics and suggested that principle on providing users with sufficient information prior to starting using the application, e.g. information about the application, on how it works, which data it uses and a contact to the application administrator for possible troubleshooting, be involved among key ethical principles.

In response to SITMP's comment that it is sometimes difficult for employees to find a suitable data source to solve a problem, or they do not even know that such an application/site exists - we implemented a new point in the guide's transparency section that organisations should promote knowledge / awareness of available data and models so that employees can find and effectively use sources already available to the organisation.

3. Identified gaps - improvement points

3.1 Content

[1] https://www.enisa.europa.eu/publications/cloud-computing-risk-assessment

As regards to content, the Ethical Code of Conduct was viewed as up to date and still relevant. This was suggested by all pilot partners.

We identified the following main three gaps / improvement points, primarily from DUET partner pilots:

A) Additions in user types to the Ethical Code of Conduct

- Athens suggested the addition of startup roles, including 'investor'/entrepreneur, to Section 3. The importance of the GDPR compliance issue with regard to private actors is stressed to raise awareness about ethics and privacy from the outset of any smart city projects.
- Regarding Section 3, Flanders suggested addition of Smart city service providers (private sector) roles and, in relation to Section 2 Flanders would like to see tips and tricks to know where the data will be stored physically and information about the regulation applied to the service provider (US, Chinese law, etc.) added to Section 2 as in their view it is a very practical question cities are coping with today.

B) Continuous monitoring of new epics

- Pilsen mentioned that the most important content to add would be the continuous monitoring of new epics and providing users with sufficient information prior to starting using the application. This should include information about the application, on how it works, which data it uses and a contact to the application administrator for possible troubleshooting, and should also be involved among key ethical principles as past epics.

C) Additions to the Preliminary considerations on ethics and the Cloud

- Athens highlighted the addition of a data-sources section defining private/public data sources and reference to ENISA to be included in the final iteration.
- Athens also mentioned that (i) a check of where the data is sourced from (private/public) - for example, check of any restrictions on private sources if you are buying them (potentially link to Cities Guide), and (ii) other materials such as Commission whitepapers (and newest documents/recommendations to review) could be added to the D1.5. According to Athens, other additional content could include guidelines on legal and ethical procedures when integrating DUET with other existing systems or newly developed systems for its further development.

D) Experts' suggestions that the ethical principles should stress privacy aspects

- Based on suggestions received during the expert review phase, the ethical principles would benefit from a stronger emphasis on the aspects of privacy. First, the ethical principles should reiterate the data minimisation principle, which is the cornerstone of any privacy-by-design approach. Second, while it is useful to suggest to decision makers that personal data should be anonymized in cases where anonymous data suffice for the intended use, the ethical principles newly recommend that the anonymisation is done at the level of the upstream data provider, who best understands the data, and before the data is supplied downstream. By only acquiring already anonymised data, the decision maker can effectively minimise sensitive data breach risks.

3.2 Structure and user experience

Athens suggested making the online website format more guided so that a user could go online and first do a viability or ethical check, in which questionnaire an answer to each question would direct the user to another section according to their answer, e.g. “What role do you have? > relevant section > will you collect personal data from citizen?”. The user would then receive a small report that indicates okay on these matters. According to Athens, this could be included on the website of DUET. For now, additional links have been provided in the current interaction to make searchability and navigation easier. However we regard it as important that the full Guide is also available so that users are alerted to issues about which they might not have been previously concerned.

Athens also pointed out to Section 2 as referring more to technical partners of DUET, rather than to the legal department (as it already has the knowledge).

Flanders highlighted that the draft D1.5 ethical principles could become part of a questionnaire directing the user depending on the answers to the next relevant topics and that the outcome could be a list of dedicated principles the city needs to check, leading to “likely ethical consideration” as mentioned in Section 3. of the D1.5.

4. Conclusion and steps to increased awareness

This document is the final in a series of deliverables concerning the Ethical Principles for using Data-Driven Decision in the Cloud". The aim of this document was to assess and improve D1.5 by highlighting the useability of the guidance and present areas of strengths and areas for improvement.

In addition to the feedback collected from DUET partners, this deliverable also presented a final version of the guidance to be published for use by other smart cities. While the final version (Annex 2) integrates the separate guides provided by deliverables D1.5, we advise any reader to use the Ethical Principles in conjunction with these deliverables because each provides a more detailed analysis of the topics at hand and their theoretical foundations.

Finally, the GSL team remains committed to increasing and spreading awareness of topics that have been analysed or in some cases merely hinted at during our work on the Cities Guide. More specifically, we aim to:

- *Sync with DUET coordinators to implement user-experience suggestions from DUET pilot cities:*
  - *Upload all legal/ethical deliverables and the final Ethical Principles on the DUET web page separately so that a city that aims to consult the guidance can navigate through url links; and*
  - *Import the checklist into an online survey format for ease of use.*
- Sync with DUET coordinators to allow for the continuous updating of epics for roles involved in data decisions in smart cities.
- *Coordinate with other DUET consortium partners with the view to (i) continue engagement with the DUET project, if the European Commission decides to prolong it, and (ii) identify new opportunities in research projects in the field*
- *Work to identify opportunities for presenting and publishing results from the guidance, with the aim to further help current and future smart cities in the implementation of ethical principles when working with data in the cloud.*

Annex 1 - Questionnaire

Questionnaire for Feedback on D1.5 Ethical Principles for using Data-Driven Decision in the Cloud

Opening questions:

1. Does D1.5 meet the following goal of providing an ethical code of conduct for cities in data-driven decision making? If so, how?
2. How have you used D1.5 to assess ethics when making data-driven decisions?
3. When implementing the DUET "ambitions / use cases", which use case you have found most challenging from the perspective of privacy or ethical considerations?
4. Please suggest any improvements in the format of D1.5, to improve the searchability and usability of output.

Section specific questions

2. Preliminary considerations on ethics and the Cloud

1. Do you consider this chapter still broadly in line/up to date and relevant for developing DUET use cases? Are there any areas that should be included beyond disruptive technologies, privacy, cybersecurity, cloud infrastructure, HPC and AI, dissemination, wider societal considerations, codes of ethics and, trust building for ethical decision making.
2. Is there sufficient information concerning the preliminary consideration on ethics and the Cloud? What other information would you like to see in this section?
3. If you have consulted this section, please describe how you utilised the guidance.

3. Duet Ethical Principles

1. Please describe how you utilised the ethical framework available in this guide. Were the steps sufficient in providing guidance?
2. Which elements do you now have a better understanding of as a result of using the ethical framework?
3. Considering the 'User Types' in the DUET Ethical Principles, which other roles could be included in this section? Currently included are the ethical considerations for the DUET Administrator, Public Servant, Policy Maker, Investor, registered user, citizens. Which other categories of information would be useful for these roles in this guidance?
4. Do you consider the ethical framework to still broadly be in line/up to date with and applicable to the developing DUET ambitions/use cases? Please consider 'The Draft Ethical Principles' on page 32 of D1.5.
5. Are any key principles missing from the 'The Draft Ethical Principles framework from your knowledge? If yes, please indicate which principles.

3.10 Privacy-by-design

1. Please describe how (if at all) you used this section in conjunction with The Cities Guide (D1.2).

4. Future work (conclusion)

1. To your knowledge, how could this section be improved with other content?

Annex 2 - “Ethical Code of Conduct” (final and consolidated version)

A. Additional considerations on ethics and the Cloud

This section further builds on the preliminary considerations which were laid out as the ethical framework for evidence-based decision making in the cloud in D1.5. The following is supplementary to Section 2, **Preliminary considerations on ethics and the Cloud,** as published in D1.5.

1. Privacy

In addition to the Privacy implications discussed in 2.2 Privacy of D1.5, the following provides extensions to the ethical considerations as requested by the pilot partners.

Ethics in public/private data sources

When selecting and possibly buying data sets, an understanding of whether a data source is based on public or private data is a key step to ensuring usage of the data does not infringe on the user’s rights. Deliverables D1.1. (Legal Landscape and Requirements Plan), and D1.2 ((Cities Guide to Legal Compliance for Data-Driven Decision Making It. 1) describe extensively the principal legal requirements in this area and provide guidance on the legal necessities that a smart city should take into account in its decision and policy making processes.

There are several areas of interest smart cities must take into account to ensure compliance when selecting a data source. These include the following concerns:

- how the data was initially collected;
- purpose limitation and why the data was initially collected and processed, including any usage limitations;
- whether the data was collected privately or is scrapped from a public database;
- anonymisation of users;
- ascertaining the consent of users and usage limits from the previous supplier;
- data purpose limitation and the danger of function creep.

When selecting a private or public data source smart cities should follow guidance to ensure compliance, as elaborated further in this deliverable (Section 3. DUET ethical principles).

Identification of the storage location of data sets

While the focus of this deliverable is European Law, it is improper to consider the protection of data subjects' personal data in smart cities without the appreciation of our connected world via data centres and storage locations. The following therefore briefly addresses how to navigate this nexus of laws to best ensure European data protection while benefiting from other international service providers.

In particular, smart cities should take steps to ensure that they have identified the location of where data will be stored, as storage is part of processing of personal data, and evaluate any potential data protection or ethical implications of this location. There are several considerations smart cities must take into account to ensure compliance when storing personal data. These include the following concerns:

- **Safeguards for European citizens including:** i) the maintenance of a European-based data centre; ii) in the case personal data must be transferred from Europe to a third country, implementing safeguards including Standard Contractual Clauses (SCCs), Corporate Binding Rules (BCRs), and ensuring a sufficient adequacy decision from the European Commission allowing the transfer of personal data to this location.
- **Mitigating risks associated with storing data:** when working with multiple databases or datasets there is a possibility personal data will be indirectly stored in locations of third countries. This raises issues when personal data is stored in a country where the government maintains a wider discretion to access personal data of companies, such as the US. If this event does arise, it is important smart cities ensure when processing personal data i) the above safeguards are in[2] place; ii) requesting notification upon the request from government authorities of the disclosure of personal information; iii) notifying users of the smart city in relevant notices and applications of the potential for their personal data to be transferred to a third country.
    - In the event that no safeguards can be attained, the smart city should seek to use an alternative data source.

B. DUET ethical principles

This is a final and consolidated version of the "Ethical Code of Conduct" (Ethics Guide).

User types

See also: User types in D1.5

The following overview contributes further user types into overarching categories and attributes to them some characteristics and potential use cases identified thus far by the DUET partner organisations. Please note that additional user types in this deliverable were not initially selected and are included as indicated on request from pilot partners. Furthermore, these roles have been included due to increased debate (among DUET and also elsewhere), that they are key stakeholder groups in the emerging smart city initiatives.

[2] Some jurisdictions have found transfers to the US not compliant with European Data Protection Law, such as France - 'CNIL concludes that transfers to the United States are currently not sufficiently regulated' https://www.cnil.fr/en/use-google-analytics-and-data-transfers-united-states-cnil-orders-website-manageroperator-comply

These user types therefore do not tie directly to DUET epics but an attempt has been made to connect them to the best fit.

Remaining user types were previously published in D1.5. These user types have been selected with regard to their likely impact on ethical considerations from the work-in-progress document on user epics as found here.

Entrepreneur/founder (private sector)

While not initially selected as a user type, Entrepreneur/Founder was requested from the pilot partner Athens to be provided in the final output.
Characteristics:

- Connecting data sources
- Restricting access to data sources
- Designing visualisation and presentation features
- Business to public information sharing

Selected epics:

- (**G15**) I want to be able to connect data sources so I can be sure that the necessary data and information is available
- (**G16**) I want to be able to restrict the access to data sources so I can be sure confidential data is not made publicly available
- **(P7**) I **want to** provide 3D data (as well as BIM data) of my envisaged major construction project to the city, thus allowing the city administration to assess my project in 3D **so I can i**nform citizens about my project in the official digital twin of the city (under the 'future' view).

Likely ethical considerations:

- Accountability and transparency - knowing and making known the origin of the data, respecting and passing on data limitations
- Data quality
- Data security
- Transparent and fair use of data models
- Privacy-by-design (GDPR compliance) and confidentiality, risks of re-identification of anonymous data
- Setting results presentations technical options (colour modes, visualisations, diagrams, etc.)

Smart city service provider roles (private sector)

While not initially selected as a user type, a role of a smart city service provider (private sector) was requested from the pilot partner Flanders to be provided in the final output.

Characteristics:

- Connecting data sources
- Restricting access to data sources
- Designing visualisation and presentation features
- Business to public information sharing

Selected epics:

- (**G39**) I want to trigger traffic model simulation by providing the possibility to lower the speed or by changing the road capacity so I can do more detailed what-if analysis
- (**G20**) I want to work in a separate environment deployed on hardware on my budget so I can be independent from the other pilots

Likely ethical considerations:

- Privacy-by-design considerations including product and organisational security and privacy
- Balancing the interests between law enforcement and keeping data separate (issue of purpose).
- Selecting trustworthy third parties to process data

Public servant / city official / city employee / urban planner (a public authority)

While included in D1.5, this user type has been extended with an additional ethical consideration indicated in bold below.

Characteristics:

- Access to data, models and results
- Setting the purpose of data processing
- Making evidence/data-based decisions
- Position to negotiate with partners (public, private) access to or sharing of data

Selected epics:

- (**G2, G3**) I can discover the causes of pollution. I can assess the impact of changes to the local situation on the traffic in my area of interest / the impact on citizens well-being in the city.
- (**P1**) I want to see all existing attributes for buildings and objects in 3D representation of the city
- (**Pilsen**) I want to connect existing data resources of the city to the digital twin and make sure they are up-to-date, interoperable, and include all available attributes, with the goal to make my daily work more efficient thanks to working with different data sources in a single environment

Likely ethical considerations:

- Finding the best data for the purpose
- Understanding the possibility of selection bias in the choice of data and what is “best”
- Do I understand the data models? Am I able to interpret their results? Am I able to use the system?
- Transparency towards citizens that their data is used, decisions made based on data concerning them
- Ensuring data is of sufficient quality to be published
- Balancing the interests between law enforcement and keeping data separate (issue of purpose).
- Asking / requesting third parties to provide data
- **Understanding through the description of the application (instructions, examples of use, information about the application, etc.) how it works, the results of analysis and to be able to decide on the correct use of the results of the analysis (application).**

DUET administrator (system administrator)

Characteristics:

- Connecting data sources
- Restricting access to data sources
- Designing visualisation and presentation features

Selected epics:

- (**G15**) I want to be able to connect data sources so I can be sure that the necessary data and information is available
- (**G16**) I want to be able to restrict the access to data sources so I can be sure confidential data is not made publicly available

Likely ethical considerations:

- Accountability and transparency - knowing and making known the origin of the data, respecting and passing on data limitations
- Data quality
- Data security
- Transparent and fair use of data models
- Privacy-by-design and confidentiality, risks of re-identification of anonymous data
- Setting results presentations technical options (colour modes, visualisations, diagrams, etc.)

Policy maker (a public authority)

Characteristics:

- Access to data, models and results
- Purpose of data processing
- Making data available as open data
- Balancing of public and private interests

Selected epics:

- (**P8**) I want to make the 3D data of the city available as open data (see data section for already opened data). The city balances the relevance of opening the data with policy objectives, the price, the relevant level of granularity and so on.
- (**P6**) I want to motivate investors of major development projects to provide 3D data during the building planning and permission process

Likely ethical considerations:

- Finding the best data for the purpose
- Understanding the possibility of selection bias in the choice of data and what is “best”
- Do I understand the data models? Am I able to interpret their results? Am I able to use the system?
- Transparency towards citizens that their data is used, decisions made based on data concerning them
- Ensuring data is of sufficient quality to be published
- Asking / requesting third parties to provide data

Investor (a private party)

Characteristics:

- Restricted / no access to confidential information

- Business to public information sharing

Selected epics:

- (**P7**) I want to provide 3D data (as well as BIM data) of my envisaged major construction project to the city, thus allowing the city administration to assess my project in 3D, so I can inform citizens about my project in the official digital twin of the city (under the 'future' view).

Likely ethical considerations:

- Ensuring data is of sufficient quality to be shared
- How can the data be further (re-)used

Registered user who can upload data (a private party)

Characteristics:

- Private to public information sharing

Likely ethical considerations:

- Ensuring data is of sufficient quality to be shared
- How can the data be further (re-)used

Citizen (a private party)

Characteristics:

- Restricted / no access to confidential information. Broad definition of confidential (incl. personal data, trade secrets, restricted use based on license information).

Selected epics:

- (**G6**) I can inspect the current traffic density.
- (**G7, G8**) I can inspect the current level of pollution.
- (**G9-G11**) I can inspect historical information.
- (**G13**, **G14**) I can provide municipalities with some data that I collect.
- (**Flanders**) I want to have an idea about the mobility flows in my city and neighbourhood.

Likely ethical considerations:

- How is my / my co-citizens' data processed?
- For what purposes?
- Is the data safe?
- Are any automated processes involved in the decision-making about me / using my data?
- Do I trust my city? Is the city transparent about its decisions and policies, and data management?

The draft ethical principles

As an updated version of the draft ethical principles released in D1.5, we have prepared a draft set of ethical principles for broader consideration among the DUET partner organisations. Below is an updated version of the principles including additional principles were derived from the following sources:

- Discussions with DUET partners;
- Extrapolation from more detailed or, conversely, high level, legal requirements and principles.

The first draft of these principles which this final version includes was derived from the following sources:

- Primary desk research scoping similar activities by smart cities elsewhere in the EU and overseas;[3]
- Discussions with DUET partners (AIV and IMEC) and external reviewers;
- Extrapolation from more detailed or, conversely, high level, legal requirements and principles.

In order to attract more detailed comments from pilot city organisations, we provided an (unexhaustive) list of broadly framed questions in order to illustrate what type of ethical considerations they should take into account. These are the questions we asked:

- *"I had been assigned a task and I was not sure whether I am allowed to use the data collected by my City for this purpose."*

- *I was not sure whether my City has data available that would help me to complete a particular task. I was unsure who to turn to to find out."*

- *"I felt that I lacked the skills and knowledge in data analytics / visualisation in order to use the available data to make a data-based decision."*

- *"I had to rely on a result given by a data processing model to carry out a task or make a decision, but I was not sure whether the result was reliable."*

- *"I wanted to use a data processing model but I was not sure that I understood fully how it works."*

- *"When working with some data (e.g. noise, pollution), I realised that the data reveal the source or the cause of these externalities (e.g., the exact location of a polluter). I was wondering if someone may be responsible for committing an administrative offence (or even a crime) and I was not sure whom to contact about this. I was not sure whether I could make this data public, and what the reaction from the general public might be.*

- *"I worked with anonymous data but after combining different data sources, they began to reveal information about identifiable persons (personal data)."*

- *"I was working with a third-party provided dataset, but realized that these are actually personal, highly sensitive data (e.g., data showing individual people's health condition, or data allowing me to learn about a person what his/her religion or sexual orientation is). I wasn't sure about whom to contact and what to do with the database."*

- *"We cooperated with a private organisation (a firm, company, non-profit) on a joint task. The private organisation planned to collect some data in the course of the task. I was not sure whether the organisation could share the data with us. I was not sure how to ask that organisation to share the data with us. The organisation refused to share the data with us and I didn't know what to do."*

- *"I wanted to present some data with help of visualisation but the system won't allow me to choose*

[3] These include, e.g., the City of Barcelona Digital Standards (https://www.barcelona.cat/digitalstandards/en/init/0.1/index.html); the Smart Flanders Open Data Charter (https://smart.flanders.be/open-data-charter/); the Toronto Quayside project (https://waterfrontoronto.ca/nbe/portal/waterfront/Home/waterfronthome/projects/quayside).

the colour I want or a type of visualisation that would work best in my opinion."

After evaluation of the comments received, the following text sets out the additional ethical principles, or guidance to principles, that further extend the ethical principles outlined in D1.5 and could be followed by DUET partner organisations in their development and execution of the DUET project, as well as other stakeholders and smart cities with similar ambitions.

1. **Accountability and data sovereignty**

   1.1. Know the origin of the data, its lawful and ethical uses, and any limitations on their sharing or publication.

   1.2. This includes understanding the origin of data when working with **private/public data sources as expressed by the European Union Agency for Cybersecurity (ENISA).**[4]

   1.3. This also includes understanding all possible locations of processing of personal data and the different data regulations this may be subject to.

2. **Transparency**

   2.1. You should know what data you collect and for what purposes.

   2.2. The data subjects (e.g. the citizens) should know what data you collect about them and for what purposes.

   2.3. Be transparent about the scope and source of the data, as well as the limitations of the data. Explain what information the data contains, how (and where) it was collected, whether it is static data, updated regularly, or real-time.

   2.4. If the data is publicly available, provide a link to the origin data repository/source url.

   2.5. Make sure that decision makers are aware of the deficiencies/limitations of the data.

   2.6. Promote knowledge of utilisable data/models within your organisation so that employees are aware that helpful data or applications may be available to carry out their tasks.

3. Data quality

   3.1. Get the best data as you can for your purposes. Best may mean:

      3.1.1. data most suited for your purpose;
      3.1.2. most complete, correct, and up-to-date data (clean data);
      3.1.3. data with a transparent track record of their collection, storage, and the log of previous processing;
      3.1.4. data with a clear licence to (further) use.

   3.2. Take active steps to ensure and maximise the quality, objectivity, usefulness, integrity and security of data.

[4] *https://www.enisa.europa.eu/publications/cloud-computing-risk-assessment*

4. Data quality for publication

4.1. If the data is sufficient for an internal use (within the services of the city), it is typically equally good for making the data publicly accessible (open).

4.2. Use open standards and open licences.

4.3. Publish / share data only after you have cleared the applicable legal requirements.

5. **Data security**

5.1. The integrity and security of data should be maximised.

5.2. Use trusted third-party services providers (e.g., approved by the future European Union Cybersecurity Certification Scheme on Cloud Services (EUCS)).

6. **Data everywhere**

6.1. Promote the use of data in public interest, be active in seeking out data that may be (re)used in public interest.

6.2. Actively explore the ways in which data can be obtained from partners (private or public) with whom you engaged in a joint activity (e.g., public procurement).

7. **Transparent and fair use of AI and computer models. Fighting the "opacity" problem.**

7.1. Cities should strive to develop the officials' ability to understand, interpret and use automated decision-making systems. They should understand at least the basics of the underlying algorithms and the data used. This can be achieved by a targeted education and training, for example.

7.2. Data subjects (citizens) should be informed about the fact that automated decisions are being taken about them and with the help of their data. To the extent possible, cities should strive to make sure that data subjects also understand the underlying algorithms, to the extent practicable.

7.3. Algorithms and automated decisions should be fair and proportional. They should not prejudice the data subjects. Even though some bias may be inherent in data, the algorithms and the data they use (or train on) should not create or perpetuate material biases (racial, ethnical, sexual, political, religious, etc.)

7.4. Ensure an element of human control over the AI:

7.4.1. Individuals to whom human oversight is assigned should fully understand the capacities and limitations of the AI system and should be able to duly monitor its operation, so that signs of anomalies, dysfunctions and unexpected performance can be detected and addressed as soon as possible.

7.4.2. Data subjects should be granted the right to appeal relating to data processing and the automated decisions that affect them.

8. **Presentation of data or results**

8.1. The way data or data-based decisions are presented should avoid creating or perpetuating bias (e.g., the use of red and green color coding for visualisations).

9. **Data ownership and management**

9.1. Data ownership typically goes hand in hand with the responsibility for data management.

9.2. Third parties contracted out for city data management should be chosen responsibly, adequate data processing agreements should be put in place.

9.3. Smart cities should understand if their data is public or private when acquiring a data set from a third party source, and the limitations on usage.

10. **Privacy-by-design**

10.1. Comply with all legal requirements when acquiring, using, or publishing personal data. (see also D1.2 Cities Guide to Legal Compliance for Data-Driven Decision Making).

10.2. If you come across a personal data breach, report it to your Data Protection Officer.

10.3. Minimise the amount of personal data obtained, used and stored.

11. **Anonymised data preference**

11.1. Do not use personal data unless it is strictly necessary for your task and proportionate to meeting the pre-defined purpose of your activity.

11.2. If anonymous data is not available, but personal data is, ensure that the data is anonymised before its further use, if possible. **Ask the upstream data provider, who best understands the data, to anonymise the data before it is supplied.**

11.3. Non-anonymised data should in no case be made public (or open data), unless strictly required for carrying out the task in question, and unless cleared by the Data Protection Officer for publication.

11.4. Where data is anonymised, do not proactively take any steps in the direction to re-identify the data (link the data to individual persons). The following techniques and procedures, for example, should be avoided unless the goal is actually to re-identify otherwise anonymous or pseudonymised data:

11.4.1. *Singling out*, which corresponds to the possibility to isolate some or all records which identify an individual in the dataset;

11.4.2. *Linkability*, which is the ability to link, at least, two records concerning the same data subject or a group of data subjects (either in the same database or in two different databases). If an attacker can establish (e.g. by means of correlation analysis) that two records are assigned to a same group of individuals but cannot single out individuals in this group, the technique provides resistance against “singling out” but not against linkability; or

11.4.3. *Inference*, which is the possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes.

11.5. If the risk of re-identification materialises on a given dataset, take all reasonable steps, seek appropriate expert advice and apply all relevant professional standards in order to mitigate

the risk of a privacy breach and further unlawful personal data processing.

12. Training and sufficient data usage information

12.1. Ensure to provide sufficient information about the application including how it works and the data the model is sourcing from.

12.2. If applicable, provide a contact to the application administrator for possible troubleshooting.

12.3. **Ensure all people involved have an understanding of these ethical principles.**