D1.3 Data Management Plan

Document outlining and monitoring BIPED's approach to dataset and data models management with regard to the FAIR principles

Project name	BIPED: Building Intelligent Positive Energy Districts
Duration	January 2024 – December 2026
Project ID	101139060
Coordinator	Technical University of Denmark
Type of action	Innovation Action
Call ID	HORIZON-MISS-2023-CIT-01
Website	https://www.bi-ped.eu/
Document name	D1.3 Data Management Plan
Document status	Final
Delivery date	28.6.2024
Dissemination	Public
Authors	Tomáš Pavelka (UTR)

Funded by the European Union

BIPED is funded under the EU Horizon Europe Research and Innovation programme. Grant ID: 101139060. Every effort has been made to ensure the accuracy of information provided in this document, which reflects only the authors' views. The European Commission is not responsible for any use that may be made of the information contained herein. Credits: cover image from Adobe Stock.

Document history

Version	Date	Contributor	Description
0.1	06.05.2024	Tomáš Pavelka (UTR)	Outline
0.2	07.06.2024	Tomáš Pavelka (UTR)	First draft
0.3	12.06.2024	Tomáš Pavelka (UTR)	Updated draft for review
0.4	24.06.2024	Shahrzad Pour (DTU), Jiří Bouchal (DRI), Blain Murphy (KPMG)	Review
0.5	26.06.2024	Vito Michele Pavese	External review
1.0	28.06.2024	Tomáš Pavelka (UTR)	Final

Table of Contents

List of Figures

List of Tables

Executive summary

The BIPED Data Management Plan (**DMP**) is a mandatory Horizont Europe M6 deliverable to ensure compliant, legal and ethical data management for the project duration and beyond.

The DMP is set up and implemented by the BIPED consortium management (and lead by a dedicated partner UTR) and is applicable to all BIPED partners involved in handling and processing data, datasets and data models for the project purposes.

The DMP does not constitute legal advice from the consortium management nor any specific BIPED partner.

The Introduction chapter sets the scene by describing the BIPED data management lifecycle and the overarching principles that the project integrates privacy-by-design and data minimisation principles to mitigate any privacy risks, and the open data and path towards maximum reusability and transfer of project results onto further stakeholders. This ties the DMP to the FAIR principles that are the backbone of an adequate DMP for Horizont Europea project purposes.

Chapter 2. Legal Framework gives an overview of the major EU level legislation applicable to the BIPED project activities related to the data (mainly the GDPR) and sets out relevant provisions of the BIPED Grant and Consortium Agreements applicable in the area.

Chapter 3. BIPED Data Management Plan sets out the DMP proper, with subchapters addressing 3.1 Overarching principles, 3.2 Data summary and typologies, 3.3 Data collection, 3.4 Data processing, 3.5 Data storage, 3.6 Data sharing (Open Access), aligning the DMP with the FAIR principles, setting out rules and guidance to mitigate any foreseeable risks related to privacy, which is also the cornerstone for project's ethical compliance. The main overarching rule that the data owner is responsible for the data management is explained on various data types and user scenarios to provide a hands-on guidance to ensure full compliance by the responsible BIPED partners.

Chapter 4. Data usage after BIPED provides initial description of aims to ensure maximum shareability of BIPED results by way of implementing the Minimum Interoperability Mechanisms (**MIMs**).

Chapter 5. Conclusions and Future Work sets out the main signposts for future updates envisaged for the DMP also with reference to the forthcoming Deliverable D1.4 Privacy and Ethical LDT Implementation Manual, which will deep dive certain select concepts already outlined by the DMP for further practical implementation, and more. The lead DMP partner UTR will also seek to use the first DMP iteration as a basis to organise a joint meeting between partners' Data Protection Officers (**DPOs**) to maximise their awareness of the BIPED project.

Annex 1 - Key definitions and terms, Annex 2 - BIPED Data and Metadata Collection Sheet template provide certain explanatory information and Annex 3 - Model Information Sheet and Consent Form (Privacy) provide a model information sheet and consent forms recommended to BIPED partners who collect personal data for BIPED purposes to use and keep on file to ensure GDPR compliance.

Introduction

The BIPED proposal stipulates that the Data Management Plan ("**DMP**") is to be a living document that presents the consortium's plan on handling data during and after the end of the project; what data will be collected, processed and/or generated, which methodology and standards will be applied (including methods like for anonymisation and pseudonymization), what will be included as Open Data and how data will be curated and preserved.

This first iteration of the BIPED DMP seeks to cover or at least outline future work related to data elements like datasets and simulation models and the impact of Artificial Intelligence (**AI**) on data creation and outcomes along with sourcing, production, collection and processing of data, their further re-use and scientific publications to disseminate BIPED results. The first DMP iteration provides the framework for project data lifecycle, privacy considerations, and the project's policies for data collection, storage, access, sharing, protection, retention, and destruction. All data will be managed in line with the FAIR principles and the Minimal Interoperability Mechanisms ("**MIMs**").

As regards data governance, the BIPED DMP also directly benefits from deliverables concerning data governance from past EU funded projects especially in the area of digital twins such as DUET (Digital Urban European Twins) H2020 project.

D1.4 Privacy and Ethical LDT Implementation Manual: Report mapping privacy and ethical policy and LDT related elements with guidelines on best practices on how to address risks. will use the DMP as a stepping stone to develop and deliver the principal mitigators to ensure legally and ethically compliant implementation of tasks specific to LDT development and functioning. Ambition is to learn lessons and develop guidance useful for other LTD initiatives and the general public.

1.1 The BIPED Data Management Lifecycle

Collection and processing:

The basis of BIPED DMP is integrating privacy-by-design and data minimisation principle at all steps of the data management lifecycle. In a project that actually has an ambition to leverage data, their combinations and data models to achieve meaningful results and benefits that are the core aims of the project, this is possible subject to certain overarching principles (in detail addressed in chapter 3.1 Overarching principles) and processes and risk mitigators to be identified and implemented at each step of the data management lifecycle, as further explained in chapters 3.2 through 3.7.

BIPED uses a mix of existing and original (new/collected/created) data to create key exploitable results subject to applicable legal standards and ethics requirements.

We envisage three overarching categories of data to be used for BIPED purposes:

a. **Original data** collected by a BIPED partner, which will be used for BIPED purposes or for dissemination of its results (e.g. data from sensors deployed by BIPED partners with view to collect and process them primarily in the BIPED context; data collected in workshops conducted by BIPED partners);

b. **Existing data already in possession** of one or more BIPED partners prior to the project's initiation, which will be used / repurposed for use in the BIPED context;

c. **Existing data sourced/procured** by a BIPED partner for the use for BIPED purposes during the project's timeline.

Examples of existing data include geospatial datasets including 2D and 3D terrain data, statistical population data, data related to energy consumptions and efficiency, simulation model results, statistical social media data, photos of locations, architectural plans, research reports.

Examples of envisaged original data: hard and soft data collected in course of workshops with local stakeholder groups.

The overarching principle is that data ownership goes hand in hand with responsibility for the data management.

Original data that will be collected and processed will be processed in line with the FAIR principles, and where they may concern personal data or mixed datasets, also in full compliance with the EU General Data Protection Regulation (**GDPR**). As explained in chapter 3.3 Data collection, the data owner for BIPED purposes is typically the "data controller" within the meaning of Art. 4(7) of the GDPR. Where data collection and processing may concern special categories of data[1], the BIPED partners shall proceed subject to strict legal requirements and in close cooperation with the consortium management.

On the basis of this first version of the DMP, the BIPED consortium management will seek to organise a joint session with Data Protection Officers (**DPOs**) from BIPED partners to establish a working group for cases involving high privacy or ethical risks.

Living document

During the Project lifecycle, the DMP will be updated based on the input from the workshops, the available data and the policy objectives. Each use case will describe what is going to be simulated, what outcomes will be produced and what original data is going to be generated in the Digital Twin and other BIPED outputs. Special consideration of vulnerable group engagement and management will be a core focus to ensure full legal (mainly GDPR) and ethical compliance.

Further re-use

By default, original data produced by the project will be licensed for reuse under CC-BY 4.0. However, more restrictive licences may apply to datasets generated out of (or in combination

[1] Art. 9 GDPR.

with) licensed private sector data and will be part of a negotiated and documented agreement.

Original data will be published along with metadata on the BIPED data and model catalogue to make it findable. The DCAT and ISO19115-based data metadata catalogue will link with the EU open data portal and connect with the EU data spaces to access and publish (meta)data.

Metadata will also be published using JSON-LD and RDFa to allow maximum findability on the web. Beyond the above standards, additional cataloguing ontologies based on Linked Data are needed to ensure that (meta)data is consistent and interoperable. This is key to information about correct use of data and simulation models related to data quality and model outputs. Additional effort is invested in multilingualism, as labelling metadata language helps support language switching and machine translation.

Data outputs will be accessible as part of the BIPED Digital Twin. Depending on the use case and the involvement of licensed data, data will be (if allowed) visible and/or downloadable. During negotiations, three policies will be negotiated:

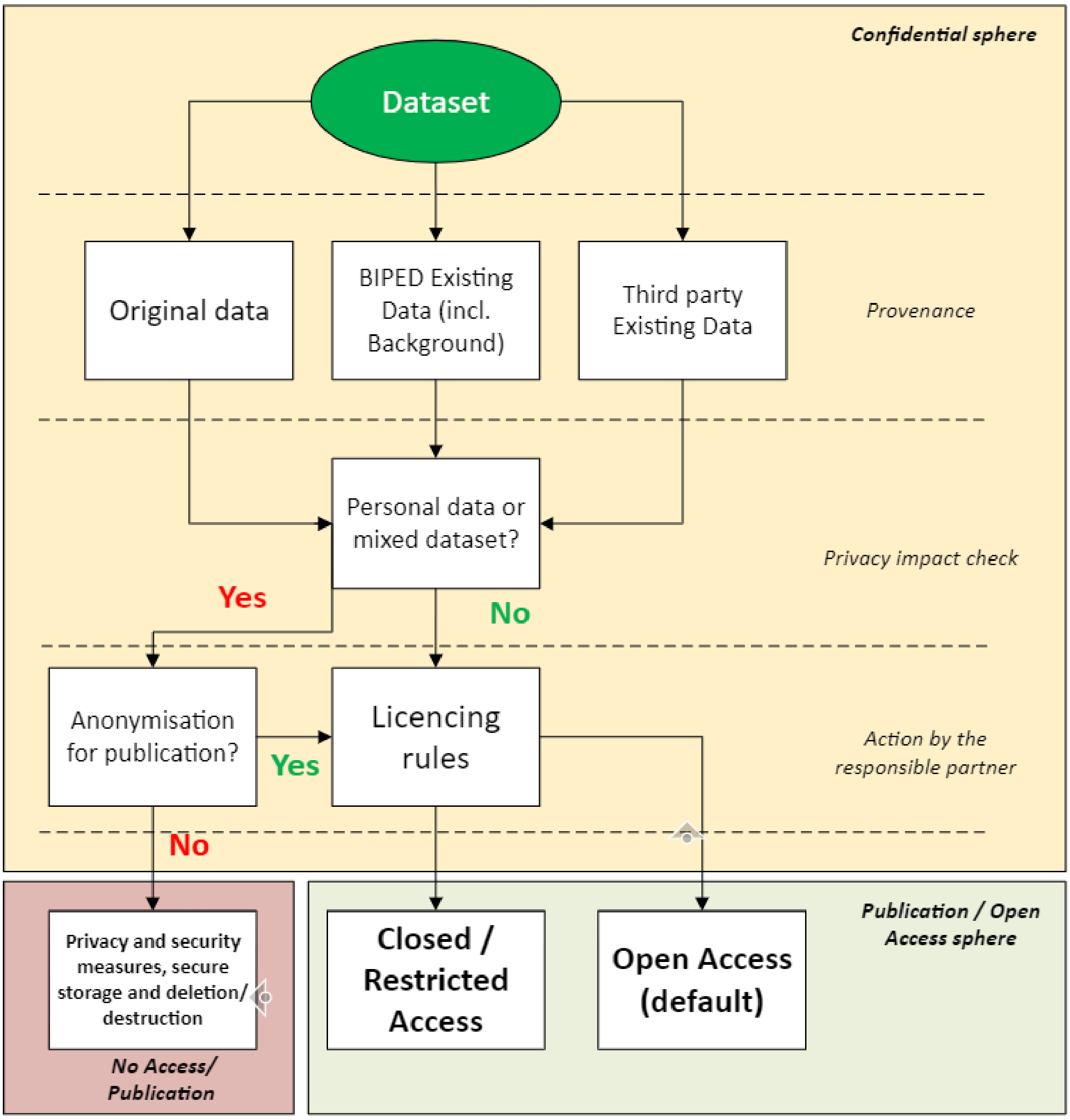
- Visibility in the Digital Twin.
- A comply or explain approach will be used to maximise Open Access during and after the use of data.
- Science outputs and journal publications will always be available under Open Access.

Different Creative Commons licences will be considered when sharing BIPED data for reusability For created open data, BIPED partners (subject to provisions of the Consortium Agreement) will use the CC-BY licence, which only requires attribution. CC-BY is in line with Open Science and Europeana track. Other data, e.g., derived from licensed data, might require more restrictive licences, such as CC-BY-ND (Attribution-No Derivatives) or CC BY-NC-ND (Attribution Non-Commercial, No Derivatives). When personal data is involved, well-described anonymisation or irreversible pseudonymisation techniques will be applied before open licensing. See in more detail chapter 3.6 Data sharing (Open Access).

Storage

BIPED data are stored on secure servers located in the EU and handled by qualified members of the Consortium under strict agreements. This ensures data access, data protection and privacy standards are in compliance with national and EU regulations. Where BIPED will be storing personal data,

Figure 1: BIPED Data Management Lifecycle


Confidential sphere
Dataset
Original data
BIPED Existing Data (incl. Background)
Third party Existing Data
Provenance
Personal data or mixed dataset?
Privacy impact check
Yes
No
Anonymisation for publication?
Yes
Licencing rules
Action by the responsible partner
No
Privacy and security measures, secure storage and deletion/ destruction
No Access/ Publication
Closed / Restricted Access
Open Access (default)
Publication / Open Access sphere

1.2 Alignment to the Principles of FAIR Data Handling

FAIR stands for Findable, Accessible, Interoperable and Re-usable, as referred to a project's research outputs – notably those made available in digital form.[2]

For an HE funded project, it is mandatory to adhere to the FAIR principles. The European Commission (2016) considers the FAIR principles fulfilled if a DMP includes the following information:

A. “The handling of research data during and after the end of the project”

B. “What data will be collected, processed and/or generated”

C. “Which methodology and standards will be applied”

D. “Whether data will be shared/made open access”, and

[2] https://force11.org/info/the-fair-data-principles/

E. “How data will be curated and preserved (including after the end of the project)”.

In the case of BIPED, this process is at the heart of this DMP set out in detail in Section 3 in the following structure:

3.2 Data summary (typologies and contents of data collected and produced)

3.3 Data collection (which procedures for collecting which data)

3.4 Data processing (which procedures for processing which data)

3.5 Data storage (data preservation and archiving during and after the project)

3.6 Data sharing (including provisions for open access)

3.7 Data usage after BIPED

2. Legal Framework

Below subchapters consider the key legislation and key provisions in the Grant and Consortium agreements governing the DMP.

Glossary of key definitions and terms used throughout this DMP, based on the above sources, is in Annex 1.

It is important to note that this document does not, and is not intended to, constitute legal advice; instead, all information, content, and materials in this document are for informational purposes only within the scope and objectives defined for the respective BIPED project deliverables. Given that this document got finalised at a certain point in time, information in this document may not constitute the most up-to-date legal or other information at the cut off date. Readers of this document and their organisations should contact their in-house team members or an attorney qualified in the concerned jurisdictions to obtain advice with respect to any particular legal matter.

2.1 The EU General Personal Data Protection Regulation (GDPR)

Regulation (EU) 2016/679 sets out the General Data Protection Regulation (GDPR) framework in the EU, notably concerning the processing of personal data belonging to EU citizens by individuals, companies or public sector/non government organisations, irrespective of their localization.

The GDPR was adopted on 27 April 2016, and became enforceable on 25 May 2018, after a two-year transition period. By then, it has replaced the previous Data Protection Directive (95/46/EC) and its national implementations. Being a regulation, not a directive, GDPR does not require Member States to pass any enabling legislation and is directly binding and applicable.

The GDPR provisions do not apply to the processing of personal data of deceased persons or of legal entities. They do not apply either to data processed by an individual for purely personal reasons or activities carried out at home, provided there is no connection to a professional or commercial activity. When an individual uses personal data outside the personal sphere, for socio-cultural or financial activities, for example, then the data protection law has to be respected.

On the other hand, the legislative definition of personal data is quite broad, as it includes any information relating to an identifiable individual, whether it relates to his or her private, professional or public life. It can be anything from a name, a home address, a photo, an email address, bank details, posts on social networking websites, medical information, or a computer's IP address.

The GDPR also defines certain special categories of personal data, the collection and processing of which has stronger legal protection.

2.2 ePrivacy Directive

Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (ePrivacy Directive) complements the EU privacy framework by providing a set of special rules protecting electronic communication data.

Insofar as relevant for this DMP, the Directive's provisions apply to a) processing of data in connection with the provision of publicly available electronic communication services, including location data of such services' users or subscribers, and to b) storing of information (data), or gaining of access to information (data) already stored, in the terminal equipment of end-users.

By way of example, activities involving collection of location data by telecom operators or access to data stored in terminal equipment by help of an app installed on users' equipment (such as mobile phone apps), or by tracking technologies processing data from terminal equipment, will typically fall within the scope of application of these rules.

It is important to note that the ePrivacy rules apply not only to personal data, but to any data in electronic communications (even though typically they will include data belonging to, or about, an individual user), and not only to data of natural persons, but to data of (or on) legal persons as well.

In terms of interplay with the GDPR, the Directive is in a relationship of speciality: provisions of the ePrivacy Directive serve to "particularise and complement" the GDPR. The GDPR, in turn, provides with respect to the ePrivacy Directive that it (GDPR) does not impose additional obligations in relation to processing of personal data in connection with the provision of publicly available electronic communication services in relation to matters for which they are subject to specific obligations with the same objective set out in the ePrivacy Directive.

The ePrivacy Directive is considered outdated and a proposal for new ePrivacy Regulation is pending in the legislative process. This may bring about important changes regarding the scope and substance of these rules. As the applicable ePrivacy rules are still only in the form of a directive, which means that the directly applicable national legislation transposing these rules must be consulted in any particular matter. There may be country-specific differences also in the way in which Member States have opted to derogate from these rules for various public policy reasons.

2.3 Regulation on free flow of non-personal data

Regulation (EU) 2018/1807 provides a legal framework for the free flow of non-personal data in the EU. Examples are machine-generated data or commercial data, which are either non-personal in nature or refer to personal data that has been made anonymous.

The regulation aims at removing data localisation requirements by individual EU Member States, make data available to competent (public) authorities for performance of their duties in accordance with the law. Examples are machine-generated data or commercial data, which are either non-personal in nature or refer to personal data that has been made anonymous.

The regulation aims at removing data localisation requirements by individual EU Member States, make data available to competent (public) authorities for performance of their duties in accordance with the law, and facilitate data porting and promote open standards, and help cooperation between authorities of different EU Member States.

This regulation governs the relationship with the GDPR in the area of mixed datasets

2.4 The PSI Directive

The directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (the "**PSI Directive**")[3] entered into force on 16 July 2019 and was due for transposition by the EU Member States by 16 July 2021.

The PSI Directive provides for an obligation of the EU Member States to make all existing documents held by 'public sector bodies' as well as public undertakings re-usable, unless access is restricted or excluded under national rules on access to documents or subject to the other exceptions. Public authorities can limit the making available of public data by imposing conditions in the standard licences as regards the re-use by the licensee dealing with issues such as liability, the protection of personal data, the proper use of documents, guaranteeing non-alteration and the acknowledgement of source.

The rules seek to stimulate the publishing of dynamic data and the uptake of Application Programme Interfaces (APIs). They will also limit the exceptions which currently allow public bodies to charge more than the marginal costs of dissemination for the re-use of their data.

The PSI Directive is based on the following core principles:

- data held by public undertakings, which the undertakings make available for re-use. Charges for the re-use of such data can be above marginal costs for dissemination;
- research data resulting from public funding – Member States are asked to develop policies for open access to publicly funded research data. New rules will also facilitate the re-usability of research data that is already contained in open repositories.
- strong transparency requirements for public–private agreements involving public sector information, avoiding exclusive arrangements.

Thanks to the PSI Directive and its modifications and implementations, the goal of making government data and information reusable has become shared at a broad European level. In addition, the awareness has been remarkably growing that as a general principle, the

[3] Available at https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32019L1024

datasets where PSI is stored must be set free by default. However, as found by the impact study, there are still certain barriers preventing the full reuse of government data and information, including data generated by the public utilities and transport sectors as well as the results from public funded R&D projects, which are key areas for BIPED Projects.

2.5 The EU AI Act

The newly adopted AI Act[4] seeks to enshrine in EU law a definition of AI systems aligned with the revised definition and contains as well a definition of General purpose models (GPAI). It applies primarily to providers and deployers putting AI systems and GPAI models into service or placing on the EU market and who have their place of establishment or who are located in the EU, as well as to deployers or providers of AI systems that are established in a third country, when the output produced by their systems is used in the EU. The Act maintains a risk-based approach and classifies AI systems into several risk categories, with different degrees of regulation applying. prohibits a wider range of AI practices as originally proposed by the Commission because of their harmful impact. The Act identifies a number of use cases in which AI systems are to be considered high risk because they can potentially create an adverse impact on people's health, safety or their fundamental rights. The Act further identifies a number of AI systems posing limited risks because of their lack of transparency (i.e. deep fakes, synthetic content) that will be subject to information and transparency requirements. The application of the AI act will be staged over two years (starting with the phasing out of the prohibited systems within six months after the act enters into force) and will require the European Commission to issue various implementing, delegated and guidelines.[5]

2.6 Grant Agreement

The following provision of the Grant Agreement (“GA”) are directly relevant for data management in the project’s lifecycle:

ARTICLE 15 — DATA PROTECTION

15.1 Data processing by the granting authority

Any personal data under the Agreement will be processed under the responsibility of the data controller of the granting authority in accordance with and for the purposes set out in the Portal Privacy Statement.

For grants where the granting authority is the European Commission, an EU regulatory or executive agency, joint undertaking or other EU body, the processing will be subject to Regulation 2018/1725.

15.2 Data processing by the beneficiaries

[4] After final approval by the Council of the EU on May 21, 2024, the EU AI Act is now set to be published in the EU's Official Journal

[5] Summary taken from https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence

The beneficiaries must process personal data under the Agreement in compliance with the applicable EU, international and national law on data protection (in particular, Regulation 2016/679 (**GDPR**). They must ensure that personal data is:

- processed lawfully, fairly and in a transparent manner in relation to the data subjects
- collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes
- adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed
- accurate and, where necessary, kept up to date
- kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data is processed and
- processed in a manner that ensures appropriate security of the data.

The beneficiaries may grant their personnel access to personal data only if it is strictly necessary for implementing, managing and monitoring the Agreement. The beneficiaries must ensure that the personnel is under a confidentiality obligation.

The beneficiaries must inform the persons whose data are transferred to the granting authority and provide them with the Portal Privacy Statement.

ARTICLE 16 — INTELLECTUAL PROPERTY RIGHTS (IPR) — BACKGROUND AND RESULTS —ACCESS RIGHTS AND RIGHTS OF USE

16.1 Background and access rights to background

The beneficiaries must give each other and the other participants access to the background identified as needed for implementing the action, subject to any specific rules in Annex 5.
‘Background’ means any data, know-how or information — whatever its form or nature (tangible or intangible), including any rights such as intellectual property rights — that is:

(a) held by the beneficiaries before they acceded to the Agreement and
(b) needed to implement the action or exploit the results.

If background is subject to rights of a third party, the beneficiary concerned must ensure that it is able to comply with its obligations under the Agreement.

16.2 Ownership of results

The granting authority does not obtain ownership of the results produced under the action.
‘Results’ means any tangible or intangible effect of the action, such as data, know-how or information, whatever its form or nature, whether or not it can be protected, as well as any rights attached to it, including intellectual property rights.

Annex 5 - Specific Rules

Open Science
Open science: open access to scientific publications

The beneficiaries must ensure open access to peer-reviewed scientific publications relating to their results. In particular, they must ensure that:

- at the latest at the time of publication, a machine-readable electronic copy of the
- published version or the final peer-reviewed manuscript accepted for publication, is
- deposited in a trusted repository for scientific publications
- immediate open access is provided to the deposited publication via the repository, under the latest available version of the Creative Commons Attribution International Public Licence (CC BY) or a licence with equivalent rights; for monographs and other long-text formats, the licence may exclude commercial uses and derivative works (e.g. CC BY-NC, CC BY-ND) and
- information is given via the repository about any research output or any other tools and instruments needed to validate the conclusions of the scientific publication.

Beneficiaries (or authors) must retain sufficient intellectual property rights to comply with the open access requirements.

Metadata of deposited publications must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: publication (author(s), title, date of publication, publication venue); Horizon Europe or Euratom funding; grant project name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action and, if possible, for their organisations and the grant. Where applicable, the metadata must include persistent identifiers for any research output or any other tools and instruments needed to validate the conclusions of the publication.
Only publication fees in full open access venues for peer-reviewed scientific publications are eligible for reimbursement.

Open science: research data management
The beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles and by taking all of the following actions:

- establish a data management plan ('DMP') (and regularly update it)as soon as possible and within the deadlines set out in the DMP, deposit the data in a trusted repository; if required in the call conditions, this repository must be federated in the EOSC in compliance with EOSC requirements
- as soon as possible and within the deadlines set out in the DMP, ensure open access — via the repository — to the deposited data, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC 0) or a licence with equivalent rights, following the principle 'as open as possible as closed as necessary', unless providing open access would in particular:
  - be against the beneficiary's legitimate interests, including regarding commercial exploitation, or
  - be contrary to any other constraints, in particular the EU competitive interests or the beneficiary's obligations under this Agreement; if open access is not provided (to some or all data), this must be justified in the DMP
- provide information via the repository about any research output or any other tools and instruments needed to re-use or validate the data.

Metadata of deposited data must be open under a Creative Common Public Domain

Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: datasets (description, date of deposit, author(s), venue and embargo); Horizon Europe or Euratom funding; grant project name, acronym and number; licensing terms; persistent identifiers for the dataset, the authors involved in the action, and, if possible, for their organisations and the grant. Where applicable, the metadata must include persistent identifiers for related publications and other research outputs.

Open science: additional practices
Where the call conditions impose additional obligations regarding open science practices, the beneficiaries must also comply with those.
Where the call conditions impose additional obligations regarding the validation of scientific publications, the beneficiaries must provide (digital or physical) access to data or other results needed for validation of the conclusions of scientific publications, to the extent that their legitimate interests or constraints are safeguarded (and unless they already provided the (open) access at publication).

Where the call conditions impose additional open science obligations in case of a public emergency, the beneficiaries must (if requested by the granting authority) immediately deposit any research output in a repository and provide open access to it under a CC BY licence, a Public Domain Dedication (CC 0) or equivalent. As an exception, if the access would be against the beneficiaries' legitimate interests, the beneficiaries must grant non-exclusive licences — under fair and reasonable conditions — to legal entities that need the research output to address the public emergency and commit to rapidly and broadly exploit the resulting products and services at fair and reasonable conditions. This provision applies up to four years after the end of the action (see Data Sheet, Point 1).

2.7 Consortium Agreement

4.4 Specific responsibilities regarding data protection
Where necessary, the Parties shall cooperate in order to enable one another to fulfil legal obligations arising under applicable data protection laws (the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and relevant national data protection law applicable to said Party) within the scope of the performance and administration of the Project and of this Consortium Agreement.

In particular, the Parties shall, where necessary, conclude a separate data processing, data sharing and/or joint controller agreement before any data processing or data sharing takes place.

5.1 No warranties
In respect of any information or materials (incl. Results and Background) supplied by one Party to another under the Project, no warranty or representation of any kind is made, given or implied as to the sufficiency or fitness for purpose nor as to the absence of any infringement of any proprietary rights of third parties.

Therefore,
− the recipient Party shall in all cases be entirely and solely liable for the use to which it puts such information and materials, and
− no Party granting Access Rights shall be liable in case of infringement of proprietary rights of a third party resulting from any other Party (or its entities under the same control) exercising its Access Rights.

8 Results

8.1 Ownership of Results

Results are owned by the Party that generates them.

8.2 Joint ownership

Joint ownership is governed by Grant Agreement Article 16.4 and its Annex 5, Section Ownership of results, with the following additions:

Unless otherwise agreed:
− each of the joint owners shall be entitled to use their jointly owned Results for non-commercial
research and teaching activities on a royalty-free basis, and without requiring the prior consent of the other joint owner(s).
− each of the joint owners shall be entitled to otherwise Exploit the jointly owned Results and to grant non-exclusive licences to third parties (without any right to sub-license), if the other joint owners are given: (a) at least 45 calendar days advance notice; and (b) fair and reasonable compensation.
The joint owners shall agree on all protection measures and the division of related cost in advance.

8.3 Transfer of Results

8.3.1
Each Party may transfer ownership of its own Results, including its share in jointly owned Results, following the procedures of the Grant Agreement Article 16.4 and its Annex 5, Section Transfer and licensing of results, sub-section “Transfer of ownership”.

8.3.2
Each Party may identify specific third parties it intends to transfer the ownership of its Results to in Attachment (3) of this Consortium Agreement. The other Parties hereby waive their right to prior notice and their right to object to such a transfer to listed third parties according to the Grant Agreement Article 16.4 and its Annex 5, Section Transfer of licensing of results, sub-section “Transfer of ownership”, 3rd paragraph.

8.3.3
The transferring Party shall, however, at the time of the transfer, inform the other Parties of such transfer and shall ensure that the rights of the other Parties under the Consortium Agreement and the Grant Agreement will not be affected by such transfer. Any addition to Attachment (3) after signature of this Consortium Agreement requires a decision of the General Assembly.

8.3.4

The Parties recognise that in the framework of a merger or an acquisition of an important part of its assets, it may be impossible under applicable EU and national laws on mergers and acquisitions for a Party to give at least 45 calendar days prior notice for the transfer as foreseen in the Grant Agreement.

8.3.5

The obligations above apply only for as long as other Parties still have - or still may request - Access Rights to the Results.

9 Access Rights

9.1 Background included

9.1.1

In Attachment 1, the Parties have identified and agreed on the Background for the Project and have also, where relevant, informed each other that Access to specific Background is subject to legal restrictions or limits.

Anything not identified in Attachment 1 shall not be the object of Access Right obligations regarding Background.

9.1.2

Any Party may add additional Background to Attachment 1 during the Project provided they give written notice to the other Parties. However, approval of the General Assembly is needed should a Party wish to modify or withdraw its Background in Attachment 1.

9.2 General Principles

9.2.1

Each Party shall implement its tasks in accordance with the Consortium Plan and shall bear sole responsibility for ensuring that its acts within the Project do not knowingly infringe third party property rights.

9.2.2

Any Access Rights granted exclude any rights to sublicense unless expressly stated otherwise.

9.2.3

Access Rights shall be free of any administrative transfer costs.

9.2.4

Access Rights are granted on a non-exclusive basis.

9.2.5

Results and Background shall be used only for the purposes for which Access Rights to it have been granted.

9.2.6

All requests for Access Rights shall be made in writing. The granting of Access Rights may be made conditional on the acceptance of specific conditions aimed at ensuring that these rights will be used only for the intended purpose and that appropriate confidentiality obligations are in place.

9.2.7

The requesting Party must show that the Access Rights are Needed.

9.3 Access Rights for implementation

Access Rights to Results and Background Needed for the performance of the own work of a Party under the Project shall be granted on a royalty-free basis, unless otherwise agreed for Background in Attachment 1.

9.4 Access Rights for Exploitation
9.4.1 Access Rights to Results
Access Rights to Results if Needed for Exploitation of a Party's own Results shall be granted on Fair and Reasonable conditions. Access rights to Results for internal research and for teaching activities shall be granted on a royalty-free basis.
9.4.2
Access Rights to Background if Needed for Exploitation of a Party's own Results, shall be granted on Fair and Reasonable conditions.
9.4.3
A request for Access Rights may be made up to twelve months after the end of the Project or, in the case of Section 9.7.2.1.2, after the termination of the requesting Party's participation in the Project.

9.8 Specific Provisions for Access Rights to Software
For the avoidance of doubt, the general provisions for Access Rights provided for in this Section 9 are applicable also to Software. Parties' Access Rights to Software do not include any right to receive source code or object code ported to a certain hardware platform or any right to receive respective Software documentation in any particular form or detail, but only as available from the Party granting the Access Rights.

3. BIPED Data Management Plan

3.1 Overarching principles

The BIPED project adheres to the following core principles of data management:

- **Data ownership goes hand in hand with an organisation's responsibility for data management.** This reflects the fact that BIPED is a consortium without legal personality. Instead, it is the individual partner organisations which must ultimately ensure and are responsible that any data they own is managed and processed lawfully.

- **Privacy by design.** When personal data is involved, it is mandatory for data controllers to implement privacy safeguards into the code, method, manner or technique of data collection and processing. These privacy safeguards serve to implement in design the core GDPR principles of data processing, such as data minimisation, or implementation of appropriate technical and organisational measures, such as pseudonymization, or encryption by default.

- **Data minimisation principle.** This principle means that only that which is necessary is used or consulted. Processing of any excess data is unnecessary, thereby creating unnecessary risks, which may vary from hacking to unreliable inferences resulting in incorrect, wrongful, and potentially dangerous decisions.[6] The European Commission also noted that "generating and processing less data limits the security risks,

[6] Mireille Hildebrandt, *Primitives of legal protection in the era of data-driven Platforms,* " Geo L. Tech. Rev. 252 (2018).

therefore the compliance with data minimisation measures also provides for security safeguards."[7] Adhering to the data minimisation principle can therefore be recommended as a good practice for handling non-personal data as well. Privacy by default concept is closely linked to the data minimisation principle, and in the GDPR terms it requires implementation of appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed.

- **Non-commercialisation of personal profiles**. Even where applicable law allows commercialisation of personal data and profiles built on such data, the BIPED consortium will refrain from any such activities. Accordingly, no personal data that cannot be anonymised will be made available by BIPED unless in very specific cases.

- **Open science:** publications, datasets and metadata are to be published/made available "as open as possible as closed as necessary", subject to appropriate privacy safeguards and licensing conditions.

- **No "high risk" artificial intelligence.** BIPED consortium and partners will not use or deploy for BIPED purposes AI systems which are classified as "high risk" under the applicable EU legislation (see further chapter Artificial Intelligence and Big Data).

3.2 Data summary and typologies

In terms of provenance, BIPED concerns three overarching types of data as follows:

a) **Original data** collected by a BIPED partner, which will be used for BIPED purposes or for dissemination of its results (e.g. data from sensors deployed by BIPED partners with view to collect and process them primarily in the BIPED context; data collected in workshops conducted by BIPED partners);

b) **Existing data already in possession** of one or more BIPED partners prior to the project's initiation, which will be used / repurposed for use in the BIPED context;

c) **Existing data sourced/procured** by a BIPED partner for the use for BIPED purposes during the project's timeline.

Examples of existing data include geospatial datasets including 2D and 3D terrain data, statistical population data, data related to energy consumptions and efficiency, simulation model results, statistical social media data, photos of locations, architectural plans, research reports.

Examples of envisaged original data: hard and soft data collected in the course of workshops with local stakeholder groups.

[7] European Commission Guidance on Apps supporting the fight against COVID 19 pandemic in relation to data protection C(2020) 2523 final.

BIPED will create an internal summary of datasets. Detailed information about the datasets used in BIPED will be collected by help of Data and Metadata Collection Sheets, template of which is provided as Annex 2.

Further definitions regarding data types are provided in Annex 1.

At this stage of the project, Deliverable 2.1 Chapter on Data Set & Data sources provides further details on the data per user scenarios and how they are connected towards BIPED objectives.

Personal data and anonymisation

Personal data is any information which is related to an identified or identifiable natural person.[8]

Examples of personal data that may be collected throughout the BIPED project include: names, physical address, email address, IP address, phone number, date of birth, employment and social background information, views and opinions including political opinions, video/audio recordings and transcripts of such recordings.

Personal data will not as a default be integrated with BIPED digital twin nor made available for publication or accessible unless they are anonymised.

Randomization and generalisation (e.g., aggregation or K-anonymity) are the main anonymization techniques. Using data processed with such techniques may significantly reduce risk of impact on individuals' privacy and of regulatory non-compliance.

The main guidance on these techniques was provided in 2014 by the Article 29 Working Party Opinion 05/2014 (the **Opinion**) in relation to the old Data Protection Directive, but is generally considered to provide useful guidance today under the GDPR regime.[9]

According to the Opinion, **Randomization** is a family of techniques that alters the veracity of the data in order to remove the strong link between the data and the individual. If the data are sufficiently uncertain then they can no longer be referred to a specific individual. Randomization by itself will not reduce the singularity of each record as each record will still be derived from a single data subject but may protect against inference attacks/risks. and can be combined with generalization techniques to provide stronger privacy guarantees.

Generalization is the second family of anonymisation techniques. This approach consists of generalizing, or diluting, the attributes of data subjects by modifying the respective scale or order of magnitude (i.e. a region rather than a city, a month rather than a week). Whilst generalization can be effective to prevent singling out, it does not allow effective anonymisation in all cases; in particular, it requires specific and sophisticated quantitative approaches to prevent linkability and inference.

The Opinion suggests that any anonymisation technique chosen should be tested based on the following three questions:

- is it still possible to single out an individual?
- is it still possible to link records relating to an individual?, and
- can information be inferred concerning an individual?

BIPED partners must note that an anonymised dataset can still present residual risks to data subjects. The Opinion identifies these major risks:

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

[8] Art. 4(1) GDPR.

[9] Article 29 Working Party Opinion 05/2014 on Anonymisation Techniques, available at https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

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

When conducting anonymisation, BIPED partners ensure that these risks are mitigated by choosing an adequate technique.

The Opinion at page 24 provides a helpful matrix with strengths and weaknesses of each anonymisation method:

Table 1 - strengths and weaknesses of anonymisation methods

Technique	Is Singling out still a risk?	Is Linkability still a risk?	Is Inference still a risk?
Pseudonymisation	Yes	Yes	Yes
Noise addition	Yes	May not	May not
Substitution	Yes	Yes	May not
Aggregation or K-anonymity	No	Yes	Yes
L-diversity	No	Yes	May not
Differential privacy	May not	May not	May not
Hashing/Tokenization	Yes	Yes	May not

Pseudonymisation consists of replacing one attribute (typically a unique attribute) in a record by another. The natural person is therefore still likely to be identified indirectly; accordingly, pseudonymisation when used alone will not result in an anonymous dataset. Nevertheless, it is discussed in this opinion because of the many misconceptions and mistakes surrounding its use. Pseudonymisation reduces the linkability of a dataset with the original identity of a data subject; as such, it is a useful security measure but not a method of anonymisation.

BIPED partners shall prefer anonymisation techniques to pseudonymisation where possible. Use of pseudonymised datasets for the digital twin or publication can be only done after sign off by respective organisations’ DPO and BIPED management team.

Confidential data and Data for publication

For BIPED purposes, irrespective of provenance, data on which the responsible BIPED partner has not yet assessed the potential privacy impact and whether they can be published must not be integrated with the BIPED digital twin nor otherwise made public/accessible outside the BIPED consortium. The following table shows this principle applied with respect to data types.

Table 2 - confidential data and data for publication

Provenance	Confidential	Anonymised and Public	Non anonymised (temporary status)
Original data produced by the BIPED partners	Raw survey/interview/sensor data Personal data of stakeholders (data from stakeholder workshops) New contacts established	Summaries of surveys/interviews Data in deliverables Contact data within deliverables	Photos/videos shot during public events Audio/videos recordings (e.g. Teams meetings) Data in internal repositories
Existing data already in possession of a BIPED partner	Data embedded in some of the Background (Annex 1 to Consortium Agreement) Contact databases	Data embedded in some of the Background solutions (Annex 1 to Consortium Agreement) Website logs and similar metrics	N/A
Existing data sourced/procured by BIPED partners for BIPED purposes	Raw data in possession of the cities/municipalities, businesses or other third parties	Free and open data (including from scientific and statistical publications)	N/A

Soft data

The European Commission / Eurostat's glossary defines soft data as data in the form of qualitative information or quantitative information resulting from an approximation of economic phenomena through surveys and polls.[10] Soft data may be nowadays derived also from other sources, such as scraping of social media and various crowdsourcing methods.

Soft data for BIPED purposes are mainly data attempting to describe and quantify perceptions or emotions of people to complement hard data models. As per the Proposal, one of BIPED's ambitions is to explore how quantitative collection of soft data can support the advancement in digital twin development, and to develop novel methodologies to gather such data quantitatively so it can be included in the digital twin.[11]

Hard data

In contrast with soft data, the European Commission/Eurostat's glossary defines hard data, or factual data, as data that refer to reliable and methodologically sound data taken from official or organisational statistics that are comparable and roughly independent from the way they were measured.[12]

For this DMP purposes, hard data that will be sourced from third parties are typically risk-free as regards privacy, but BIPED partners are asked to run a basic sanity check on any privacy implications, fully comply with applicable licencing requirements and consider ethical implications of using the datasets for BIPED purposes especially when combining the data with other datasets. Details are given in chapter 3.3 Data collection below.

[10] Available at https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Soft_data

[11] Part B - 3.

[12] Available at https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Hard_data#:~:text=Hard%20data%2C%20also%20called%20factual,the%20way%20they%20were%20measured.

3.3 Data collection

Data collection is considered by legislation as a type of data processing operation[13], and thus needs to be fully in line with legal requirements on data processing.

The requirements for collection depend on the provenance of the data (original/existing data):

Original data

Collecting original data for BIPED purposes will be subject to GDPR requirements if concerning personal data or mixed datasets (and special categories of personal data), and general legal requirements if concerning non-personal data. In both cases it is subject to ethical requirements and considerations, to be further addressed by the Deliverable D1.4 Privacy and Ethical LTD Implementation Manual.

The BIPED partner that has collected the data is the data owner for BIPED purposes, **i.e. the entity ultimately responsible for the data governance regarding the dataset**. In GDPR terms this is typically the "data controller" and is subject to obligations under the GDPR[14]. However, the role of data owner/controller may shift to a different BIPED partner if such partner assumes an active role in determining the purposes and means of the processing. This may be the case if a BIPED partner is a KPI owner where such KPI requires processing of dataset collected/created earlier by another BIPED partner. In another cases, however, one BIPED partner may be considered merely "processing"[15] the data on behalf of the BIPED partner which remains the "controller", if the latter still determines the purpose and means of processing. Finally, there may be cases where two or more BIPED partners are considered "controllers" of a single dataset. These situations must be assessed on a case by case basis.

Partners shall in each such case agree on a clear assignment of data ownership and responsibilities before they begin processing the data, for example integrating the data into the BIPED platforms or combining them with other datasets.

In case of personal data management and where appropriate, BIPED partners will conclude a separate data processing, data sharing and/or joint controller agreement as per clause 4.4 of the Consortium Agreement.

All partners and parties must comply with applicable licensing conditions where they are in the position of a licensee.

Existing data already in possession of a BIPED partner

Responsibility and compliance with any and all legal requirements for providing and using such data for BIPED purposes is the responsibility of the BIPED partner providing that data, unless specifically agreed with another BIPED partner (for example, the relevant KPI owner) that the other partner will assume ownership/determine the purpose and means of processing the relevant data.

[13] Art. 4(2) GDPR.

[14] Art. 24 et seq. GDPR.

[15] In case of personal data the obligations of the "processors" are set out by Art. 28 GDPR.

Per clause 16.1 of the Grant Agreement (Background and access rights to background), the *beneficiaries must give each other and the other participants access to the background identified as needed for implementing the action, subject to any specific rules in Annex 5.*
'Background' means any data, know-how or information — whatever its form or nature (tangible or intangible), including any rights such as intellectual property rights — that is:

(a) held by the beneficiaries before they acceded to the Agreement and
(b) needed to implement the action or exploit the results.

If background is subject to rights of a third party, the beneficiary concerned must ensure that it is able to comply with its obligations under the Agreement.

In case of personal data management, BIPED partners will conclude a separate data processing, data sharing and/or joint controller agreement as per clause 4.4 of the Consortium Agreement.

All partners and parties must comply with applicable licensing conditions where they are in the position of a licensee.

Existing data sourced/procured from an upstream provider

Whether personal or non-personal data, third-party provided data may carry less privacy impact risk and other risks for BIPED, because the upstream third party that is sharing or selling the data is primarily responsible for legal compliance of the dataset. The third-party provider typically also specifies the licensing conditions and the purposes for which the data may further legally be used.

Third-party data may only be used with a sufficient legal basis:

- data purchased under an appropriate commercial user licence
- data provided/obtained for free based on an individually negotiated licence/access
- data generally available for free under open access licence
- in case of personal data, the provider has a legal basis to share (transfer) the data downstream and the recipient has a legal basis to process them

The BIPED partner that has purchased/obtained the third party data is considered the data owner and responsible for compliance with any and all legal requirements for obtaining and using such data for BIPED purposes, unless specifically agreed with another BIPED partner (for example, the relevant KPI owner) that the other partner will assume ownership/determine the purpose and means of processing the relevant data.

In case of personal data management and where appropriate, BIPED partners will conclude a separate data processing, data sharing and/or joint controller agreement as per clause 4.4 of the Consortium Agreement.

All partners and parties must comply with applicable licensing conditions where they are in the position of a licensee.

Soft data

Soft data are typically collected via surveys but increasingly also with help of large-scale or crowdsourcing methods. Information from e.g. sensors in mobile phones (wi-fi sniffers) or images from car dash cameras can be useful to offer novel perspectives and insights into complex urban environments on how urban spaces are being used and on perceptions of the environment (which has been proven having significant impact on activities), hence, for BIPED purposes, on the energy profile of a district.

Data management of soft data is closely linked to the way they are collected (for example, surveys).

One of BIPED's ambitions is to develop and implement a methodology to collect soft data on a large scale and in a quantitative manner (on perception of the district by its population) to be integrated to the BIPED digital twin. Soft data that aim to capture properties of perceptions, emotions and feelings of people need to be often collected based on personal data and personal profiles.

Where quantitative soft data will be sourced from pre-existing city-level or 3rd party sources, it is reasonable to expect that such data will be fully curated at the upstream level. **However, responsible BIPED partners will seek to verify that the data is anonymised (statistical data) before integrating the data into the BIPED digital twin.**

Where personal data as a basis for soft data is collected (created) by BIPED partners, they ensure this is done based on:

- data subject's consent[16] (using Model Informed Consent Form in Annex 3 below or equivalent compliant consent form), or
- reason that collecting such data is necessary for performance of a task carried out in the public interest or in the exercise of official authority vested in the controller[17] (in case of BIPED partners - public authorities for which it is applicable),
- legitimate interest legal basis[18] and strictly for non-commercial purposes.

In case of any large-scale collection and processing of personal data to create soft data (such as web scraping), the responsible BIPED partner will ensure that this is done in compliance with GDPR requirements, possible licencing/use restrictions of the target platforms/websites, and subject to a Data Protection Impact Assessment (**DPIA**) where necessary.[19] BIPED partners consult their Data Protection Officers (**DPOs**) to ensure these requirements are met.

Collection of personal data to create soft data, which may include **special categories of personal data** (including data revealing political opinions and religious or philosophical beliefs) may be only collected based on:

- data subject's explicit consent[20] (using Model Informed Consent Form in Annex 3 below or equivalent compliant consent form),
- the fact that data subject made the personal data manifestly public,[21]
- data is necessary for scientific research or statistical purposes.[22]

Responsible BIPED partners consult their Data Protection Officers (**DPOs**) to ensure these requirements are met.

Table 3: data collection according to provenance and confidential/anonymised status

Provenance	**Confidential**	**Anonymised and Public**	**Non anonymised (temporary status)**

[16] Art. 6(1)(a) GDPR.
[17] Art. 6(1)(e) GDPR.
[18] Art. 6(1)(f) GDPR.
[19] Art. 35 GDPR.
[20] Art. 9(2)(a) GDPR.
[21] Art. 9(2)(e) GDPR.
[22] Art. 9(2)(j) and Art. 89 GDPR.

Original data produced by the BIPED partners	Surveys Interviews Sensor data Soft data sources	Newsletters Publications Open Access repositories	Workshops Interviews Internal repositories
Existing data already in possession of a BIPED partner	Access to Background and Results (as per Consortium Agreement) Seamless access by BIPED partners	Seamless access and use during project execution	N/A
Existing data sourced/procured by BIPED partners for BIPED purposes	Licensed access and use during project execution	Free and open access and use during project execution	N/A

The following table gives a list of prioritised datasets (as set out by Deliverable 2.1) indicating preliminarily the “data owners” for purposes of this DMP version (subject to further adjustments during the project lifecycle).

Table 4: prioritised datasets with preliminary indication of data owners

Group	**Variable Name**	**Details**	**Data owner (BIPED partner)**
Digital Twin	Building footprints		VCS
Digital Twin	3D city model	only Brabrand so far	VCS
Digital Twin	Orthophotos spring Webmercator	web mercator CRS for easier web integration	VCS
Digital Twin	Terrain and surface raw data	resolution 0.4m	VCS
Cross-sectoral: Environmental	weather data: temperature, wind, precipitation		DTU
Cross-sectoral: Social	Demographics: Population age distribution	on household level	DTU
Cross-sectoral: Social	Demographics: Population gender distribution	on household level	DTU
Cross-sectoral: Social	Demographics: Population ethnicity distribution	on household level	DTU
Cross-sectoral: Environmental	Land Use pattern	on parcel level	DTU
Cross-sectoral: Environmental	green space		DTU
Cross-sectoral: Environmental	air quality model	continuous spatial dataset covering the AOI	UWB
Cross-sectoral: Environmental	road network / segments	road segments input data for space syntax analysis (connectivity / accessibility	DTU, AIT, INNO

Mobility	Road network		RT
Mobility	Traffic generators (zones)		RT
Mobility	OD matrix	OD matrix connecting traffic generators and network	RT
Mobility	Sensor traffic data		INNO, RT
Mobility	Sensor data Calibration of traffic model		RT
Cross-sectoral: Environmental	Noise measurements	on sensor level	UWB
Cross-sectoral: Environmental	Noise model	continuous spatial dataset covering the AOI	UWB
Energy	Enriched Dataset on District heating from Kredsløb	Demand from individual buildings in Brabrand	CDK
Energy	Power Grid: Consumption data based on user		CDK, DTU
Energy	Power Grid: Dataset of substations		CDK, DTU
Energy	Power Grid: Aggregate consumption data on each building		CDK, DTU
Cross-sectoral: Environmental	Altitude model		DTU
Energy	District heating circuit and infrastructure	Location of district heating circuit lines and distribution system	DTU
Energy	Energy infrastructure	Location of power infrastructure and distribution system	DTU
Meteorological forecasts	Energy system	Standard MET forecasts	(unclear)
Cross-sectoral: Environmental	Identified infrastructure resilience options on energy systems		DTU

As regards possible personal data collection, the potential for collection of this data type is in WP3 for PED stakeholder engagement. Deliverable D3.1 BIPED Community provides an overview of tools and framework for stakeholder engagement and the consortium management will seek a tight alignment with the responsible partners (AAKS) to ensure full GDPR compliance of these activities.

3.4 Data processing

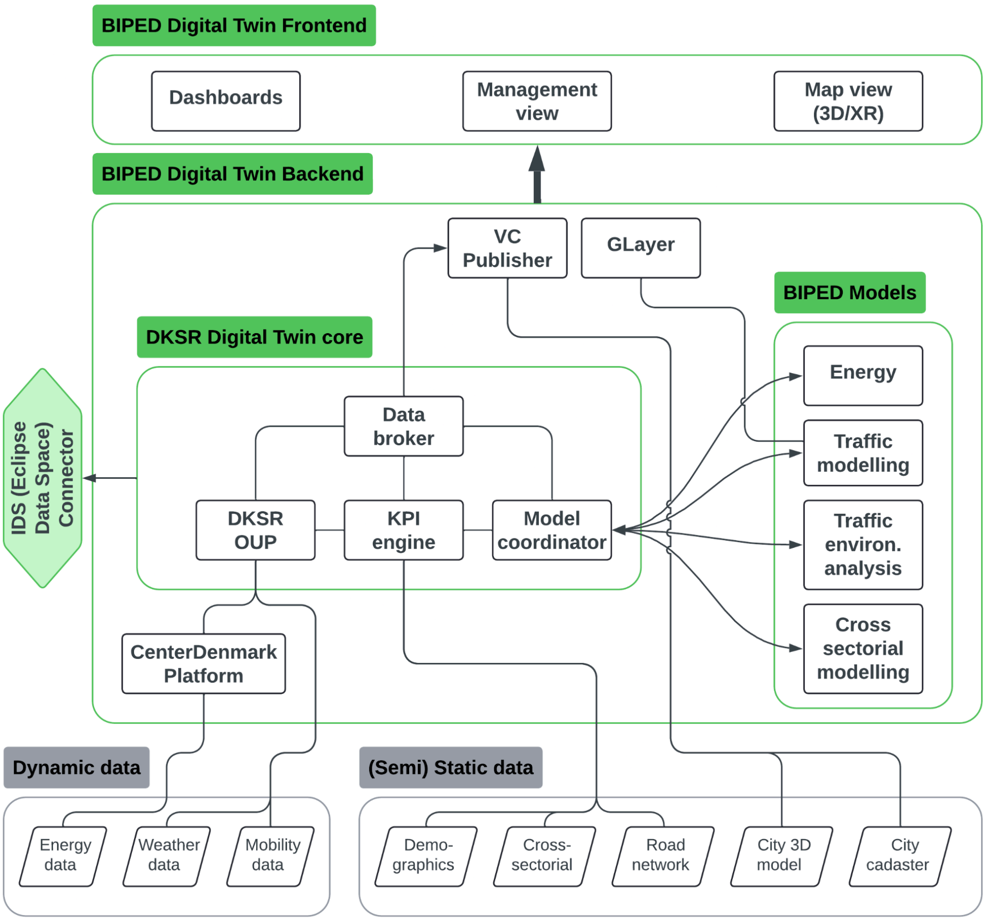
Data processing is in short any operation run on or with data, including: collection, recording, obtaining data from a third party (downloading), organisation and structuring of data, storage, adaptation or alteration, consultation and use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction of the data.

For more detail on these terms, see Annex 1.

As regards responsibility for compliance with legal requirements in data processing operations, the principles described in chapter 3.3 on data collection apply depending on the provenance and licensing restrictions of the data.

The BIPED utilises a layered architecture for processing and storing of data using three distinct components, which is shown in the Figure below. At the bottom are the data sources, these are in grey to indicate that they are exogenous to the system and not the product of this work. In the middle, the BIPED Digital Twin Backend includes the components and modules to ingest, stream, store, model and simulate the data in a consistent and structured way.

Figure 2: BIPED layered architecture



Deliverable 2.1 provides further details on these components and ingestion paths in chapter 3.

Personal data

Personal data of any provenance must be first anonymised before they can be integrated into the BIPED digital twin or otherwise processed and/or published.

The data owner is responsible for selection and deletion / destruction of any personal data and materials containing such data in the event that storing that data is no longer necessary (in line with the data minimisation principle).

The BIPED partners who are data owners ("data controllers" in GDPR terms) must ensure that they maintain record of all processing activities under their responsibility. Per Art. 30 of the GDPR, the record shall contain the following information:

(a) the name and contact details of the controller and, where applicable, the joint controller, the controller's representative and the data protection officer;
(b) the purposes of the processing;
(c) a description of the categories of data subjects and of the categories of personal data;
(d) the categories of recipients to whom the personal data have been or will be disclosed including recipients in third countries or international organisations;
(e) where applicable, transfers of personal data to a third country or an international organisation, including the identification of that third country or international organisation and, in the case of transfers referred to in the second subparagraph of Article 49(1), the documentation of suitable safeguards;
(f) where possible, the envisaged time limits for erasure of the different categories of data;
(g) where possible, a general description of the technical and organisational security measures referred to in Article 32(1) of the GDPR.

BIPED partners who are personal data controllers also ensure that they apply adequate security measures (Art. 32 of the GDPR); anonymisation and erasure steps are covered in Chapter 3.2 Data summary and typologies.

In case of any data breach, BIPED partners who are personal data controllers follow the notification obligations under Art. 33 and 34 of the GDPR subject to timelimits provided therein.

Soft data

One of the BIPED objectives is to promote integration of soft data into other datasets within the digital twin.

From a data management perspective, this may be done essentially at two stages:

- combination of soft data with hard data at the input stage, or
- combination of the results provided by processing operations on two distinct datasets / using distinct data models.

Personal data that may serve as a source for deriving soft data are subject to rules on collection/processing described in chapters 3.3 Data collection and 3.4 Data processing.

Ethical aspects of soft data processing will be deep-dived in the upcoming deliverable D1.4 Privacy and Ethical LDT Implementation Manual.

Artificial Intelligence

Even though the European Commission AI Watch working documents define AI fairly broadly and propose several definitions,[23] the newly adopted AI Act defines the AI system as *"a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments*."

Given the inclusion of elements of operational autonomy and adaptiveness after deployment, no BIPED tool/model is likely to fall in the scope of an AI system as per the AI Act definition. BIPED will in any event not deploy any artificial intelligence (AI) systems falling into the "high risk" category as per the approved text of the EU AI Act.[24] Even though BIPED results may be informative various stakeholders involved in management and operation of critical digital infrastructure, road traffic or in the supply of water, gas, heating or electricity, the consortium is of the view the digital twin or its uses would not normally qualify as "safety components" in such management and operation and thus fall out of scope of the AI Act. Recital 55 of the AI Act explains that *Safety components of critical infrastructure, including critical digital infrastructure, are systems used to directly protect the physical integrity of critical infrastructure or the health and safety of persons and property but which are not necessary in order for the system to function. The failure or malfunctioning of such components might directly lead to risks to the physical integrity of critical infrastructure and thus to risks to health and safety of persons and property. Components intended to be used solely for cybersecurity purposes should not qualify as safety components.* As a preliminary conclusion therefore, the risk of any breach of the newly adopted AI Act (when it becomes effective) in the BIPED context is low.

The following table shows the main processing operations to be run on data for BIPED purposes.

Table 5: processing of data according to provenance and confidential/anonymised status

Provenance	Confidential	Anonymised and Public	Non anonymised (temporary status)
Original data produced by the BIPED partners	Anonymisation Statistical evaluation Metadata generation Visualisation	Statistical evaluation Visualisation/publication via accessible models/demonstrators Analytics	Selection/deletion/destruction Blurring of identities
Existing data already in possession of a BIPED partner	Anonymisation Statistical evaluation Metadata generation	Statistical evaluation Visualisation Analytics	N/A

[23] AI Watch. Defining Artificial Intelligence 2.0, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC126426/jrc126426_ai_watch_defining_artificial_intelligence_2.0_final_29-10-2021.pdf

[24] REGULATION (EU) 2024/…OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of …laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

Existing data sourced/procured by BIPED partners for BIPED purposes	Anonymisation Statistical evaluation Metadata generation	Statistical evaluation Visualisation/publication via accessible models/demonstrators Analytics	N/A

3.5 Data storage

Google Drive is the selected tool as BIPED's data and information repository. This includes both the project deliverables (including relevant references utilised for their production or generated from them as project publications, e.g. journal articles, conference papers, e-books, manuals, guidelines, policy briefs etc.) and any other related information, including relevant datasets, with the exceptions specified here below.

Google Workspace (formerly G-suite) and Google Cloud Platform services are certified under the newly re-established EU-US Privacy Shield Framework.[25] When using these services, users' data are stored on servers in the US or in the EU and are predominantly processed by Google in its role as a data processor.[26] The use of these services may be considered GDPR-compliant subject to adherence to further risk mitigating measures. BIPED has implemented the following additional mitigators:

- Access to the drives where BIPED data is stored is limited to selected individuals from BIPED consortium partners, which are members of the project teams and are added to the access list on a "need to know" basis only. Mutual access to data between BIPED partners is governed by the Consortium Agreement and summarised in chapter 3.3 Data collection.

- Where access is provided to a third-party individuals or organisations (outside of BIPED consortium), only "need to know" access will provided to BIPED drives and always on a restricted basis to information / materials that are necessary for the third party involvement in the project.

- Raw data and information containing personal data and with risk of containing special categories of personal data per Art. 9 of the GDPR, such as video/audio recordings from the stakeholder workshops, will be stored and processed on secured servers of the responsible BIPED partner organisations away from the BIPED Google drives. Such data/information may be only uploaded to BIPED drives, integrated to the digital twin or otherwise made public only after being anonymised by the responsible BIPED partner, in line with principles set out by this DMP above.

[25] COMMISSION IMPLEMENTING DECISION of 10.7.2023 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate level of protection of personal data under the EU-US Data Privacy Framework

[26] https://policies.google.com/privacy

Until the responsible BIPED partner establishes a final dataset or result for publication, its **intermediate** versions are deemed confidential and handled according to principles set out in respective chapters of this DMP above.

Table 6: storage of data according to provenance and confidential/anonymised status

Provenance	Confidential	Anonymised and Public	Non anonymised (temporary status)
Original data produced by the BIPED partners	Individual partner repositories Common project repository	Project website Open access repository	Individual partner repositories Common project repository
Existing data already in possession of a BIPED partner	Specific software repositories Individual partner repositories	Project website Open access repository	N/A
Existing data sourced/procured by BIPED partners for BIPED purposes	Individual partner repositories Third party repositories Cloud repositories	Open access repositories Cloud repositories Third party repositories	N/A

3.6 Data sharing (Open Access)

BIPED Open Access Principles

BIPED results will be published and disseminated by (for example) publishing practical guidelines, reusable models, algorithms, data models, components, training materials and deliverables.

These will be available principally via the BIPED toolkit and BIPED website. BIPED will also seek to utilize the European Open Science Cloud and other platforms (OpenAIRE , Open Research, Zenodo) where appropriate.

As per the Grant Agreement and following the applicable HE rules, the following categories of data and results will be made available in the following manner:

- **written project deliverables** in readable PDF format on BIPED website (trusted repository) under CC BY or CC BY-NC / CC BY/ND license;
- deposited **datasets/models** will be available / downloadable from BIPED toolkit and/or BIPED website under the CC BY or CC0 license under the principle "*as open as possible as closed as necessary*",
- **metadata** of deposited publications/datasets will be made available in a machine-readable and actionable format under CC0 license in line with the FAIR principles. Metadata will be published under the DCAT standard[27] and other standards (SO19115-based data metadata catalog will link with the EU open data portal and connect with the EU data spaces to access and publish (meta)data; and JSON[28]-LD and RDFa to allow maximum findability on the web.

[27] https://www.w3.org/TR/vocab-dcat-3/

[28] JavaScript Object Notation, a simple but powerful format for data. It can describe complex data structures, is highly machine-readable as well as reasonably human-readable, and is independent of platform and programming language, and is therefore a popular format for data interchange between programs and systems. From https://opendatahandbook.org/glossary/en/terms/json/

BIPED will respect the requirement of HE rules/Grant Agreement to make available publications, data and metadata "*as soon as possible*". The general deadline for publication will be **6 months**, which the responsible BIPED partners typically require to finalize discussions with the Consortium management and, where applicable, third parties, regarding the scope of the appropriate licenses and any other technical questions regarding the publication /release. There may be justified exceptions to this general deadline, for example, as required by the submission process and deadlines for publication of journal articles that are based on a specific deliverable.

Background information

Any Background information (as defined by Article 16.1 of the Grant Agreement), when provided for BIPED purposes, is treated in line with rules in chapter Existing data already in possession of a BIPED partner. Limits and restrictions on Background information are set out in Annex 1 to the Consortium Agreement and overrule the principle of open access where necessary.

Ownership of results

Per clause 8 Consortium Agreement, results are owned by the Party (i.e. a BIPED partner) that generates them. Cases of joint ownership are governed pursuant to Grant Agreement Article 16.4 and Annex 5 and clause 8.2 of the Consortium Agreement.

BIPED parties coordinate on these issues in the spirit of the Consortium Agreement and any disputes will be settled according to mechanisms set out therein.

Table 7: sharing of data according to provenance and confidential/anonymised status

Provenance	Confidential	Anonymised and Public	Non anonymised (temporary status)
Original data produced by the BIPED partners	Personal email communication Shared repositories	Project website Open access repository BIPED Toolkit	N/A
Existing data already in possession of a BIPED partner	Personal email communication Shared access to software repositories	Project website Open access repository BIPED Toolkit	N/A
Existing data sourced/procured by BIPED partners for BIPED purposes	Personal email communication Shared repositories Access to third party repositories	Project website Open access repository BIPED Toolkit	N/A

4. Data usage after BIPED

MIMs: one of BIPED's ambitions is to utilize Minimal Interoperability Mechanisms (**MIMs**), which are interoperability standard to enable a minimal but sufficient level of interoperability for data, systems, and services specifically in the context of smart city solutions.[29]

BIPED will use the existing efforts on MIMs led by OASC to support interoperability between these different aspects of data use to better develop LDT solutions for this project. In particular, a key objective of the MIM1 on Context Information is to support the comprehensive and integrated use, reuse and sharing of data, enabling the bringing together

[29] https://oascities.org/minimal-interoperability-mechanisms/

of context information from different systems and sources through a web-based API, thus turning data into a strategic resource, which is also at the heart of BIPED's use of data for LDT development. Similarly, MIM2 on Data Models, which identifies interoperability mechanisms to support cities and communities to use consistent and machine-understandable definitions of all entities for data, can support a smoother integration of data from different sectors by BIPED. In addition, MIM7 on geospatial data, which supports the interoperability to integrate and transfer data between internal and external IT systems, is also highly relevant to BIPED's LDT solutions. Other MIMs, such as those on security and personal data management, may also be useful and require further assessment. As most of the MIMs are still under development, BIPED will use existing progress on MIMs to support its work, but will also use progress made by BIPED to support further development of the MIMs. This process will also provide a good basis for later consideration of the wider applicability and replicability of the BIPED solution beyond the pilot settings. Building on the launch of the first release of the Digital Twin Platform & Architecture (D2.1), the next step in the coming months will be to specify a plan for how MIMs will be used in the further development of the BIPED platforms.

5. Conclusions and Future Work

The following tasks are envisaged in M6-12 of the project:

- Dedicated legal/ethics partner (UTR) to initiate a joint call/meeting of Data Protection Officers (**DPOs**) of BIPED partners to discuss the initial version of this DMP, update them on the project status and and seek to obtain their comments.
- Updates to this DMP to be aligned with relevant project milestones to reflect progress of BIPED partners in collection and procurement of data for BIPED digital twin, integrate experience from workshops and other field work with stakeholders, and reflect the more crystalized view on how BIPED results will be shared and accessible after the project's execution (i.e. beyond M36).
- The DMP to be updated with details of processes on how BIPED will achieve application of various standards to make data/sets and results open and available in line with the project ambition and FAIR standards. This very much depends on what data will be ultimately used for the digital twin and BIPED results so further specification will become possible at a later stage.
- The DMP will be updated (if necessary) to fully align with the upcoming Deliverable D1.4 on Privacy and Ethical LDT Implementation Manual.

5. References

European Data Portal: Open Data Goldbook for Managers and Data Holders (2018) https://data.europa.eu/sites/default/files/european_data_portal_-_open_data_goldbook.pdf

Eurostat Glossaries
https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Thematic_glossaries

EU AI Act approved regulation text https://artificialintelligenceact.eu

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/eli/reg/2016/679/oj

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32002L0058

Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union https://eur-lex.europa.eu/eli/reg/2018/1807/oj

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast) https://eur-lex.europa.eu/eli/dir/2019/1024/oj

Summary of AI Act contents https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence

Article 29 Working Party Opinion 05/2014 on Anonymisation Techniques, available at https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

AI Watch. Defining Artificial Intelligence 2.0, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC126426/jrc126426_ai_watch_defining_artificial_intelligence_2.0_final_29-10-2021.pdf

COMMISSION IMPLEMENTING DECISION of 10.7.2023 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate level of protection of personal data under the EU-US Data Privacy Framework https://eur-lex.europa.eu/eli/dec_impl/2023/1795/oj

DCAT Data catalogue https://www.w3.org/TR/vocab-dcat-3/

JavaScript Object Notation https://opendatahandbook.org/glossary/en/terms/json/

OASC Cities MIMs initiative https://oascities.org/minimal-interoperability-mechanisms/

Annex 1 - Key definitions and terms

Table 1: Data according to protection

Category	Definition	Comment/example
Personal data	Any information relating to an identified or identifiable natural person (data subject); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[30].	**Examples**: name and surname; a home address; an email address such as name.surname@company.com; an identification card number; location data (for example the location data function on a mobile phone); an Internet Protocol (IP) address; a cookie ID; the advertising identifier of a mobile phone; data held by a hospital or doctor, which could be a symbol that uniquely identifies a person. Conversely, personal data are not (for example): a company registration number; an email address such as info@company.com; anonymised data. **Comment**: the question of whether data relate to a certain person is something that has to be answered for each specific data item on its own merits[31].
Special categories of personal data	Personal data directly or indirectly revealing: racial or ethnic origin; political opinions; religious or philosophical beliefs; trade union membership; genetic data; biometric data; data concerning health; data concerning a natural person's sex life or sexual orientation[32].	"Genetic data" means personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person in question; "biometric data" means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data; "data concerning health" means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status[33].
Non-personal data	Data other than personal data.	Guidance defines these data by origin either as: data which originally did not relate to an identified or identifiable natural person, such as data on weather conditions and air pollution generated by sensors installed on wind turbines or data on maintenance needs for industrial machines; and data which were

[30] Article 4(1) GDPR.
[31] WP29 Opinion 4/2007 on the concept of personal data.
[32] Article 9(1) GDPR.
[33] Article 4 GDPR.

		initially personal data, but were later made anonymous[34].
Mixed dataset	Dataset or a model that contains at least one personal data point.	**Comment**: Mixed datasets represent the majority of datasets used in the data economy and are common because of technological developments such as the Internet of Things (i.e. digitally connecting objects), artificial intelligence and technologies enabling big data analytics.[35] GDPR must be observed for the personal data part of the set[36]. **Example**: data related to the Internet of Things, where some of the data allow assumptions to be made about identifiable individuals (e.g. presence at a particular address and usage patterns).
Mixed dataset with inextricably linked personal and non-personal data	Situation whereby a dataset contains personal data as well as non-personal data and separating the two would either be impossible or considered by the controller to be economically inefficient or not technically feasible[37].	**Comment:** if at least one personal data point is inextricably linked to the non-personal data in a given mixed dataset, the whole dataset falls under the definition of "personal data"[38]. Separating the dataset may decrease the value of the dataset significantly. In addition, the changing nature of data (dynamic data) makes it more difficult to clearly differentiate and thus separate between different categories of data. In practice, mixed datasets will generally be considered personal data[39]. **Example:** when buying CRM and sales reporting systems, the company would have to duplicate its cost on software by purchasing separate software for CRM (personal data) and sales reporting systems (aggregated/non-personal data) based on the CRM data.

[34] Communication from the Commission to the European Parliament and the Council - Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union - COM(2019)250.

[35] Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union; COM(2019) 250 final, page 8.

[36] *Commission Staff Working Document, Impact Assessment accompanying the document Proposal for a Regulation of the European Parliament and of the Council on a framework for the free flow of non-personal data in the European Union* (SWD(2017) 304 final), part 1/2, p. 3, 'regardless of how much of personal data are included in mixed datasets, GDPR [the General Data Protection Regulation] needs to be fully complied with in respect to the personal data part of the set.

[37] Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union; COM(2019) 250 final, page 10.

[38] Article 2(2) of the Free Flow of Non-Personal Data Regulation: "In the case of a dataset composed of both personal and non-personal data, this Regulation applies to the non-personal data part of the dataset. Where personal and non-personal data in a dataset are inextricably linked, this Regulation shall not prejudice the application of Regulation (EU) 2016/679."

[39] Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union; COM(2019) 250 final, page 10.

Trade secrets	Information which meets all of the following requirements: (a) it is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question; (b) it has commercial value because it is secret; (c) it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret[40].	**Examples**: undisclosed know-how and business or technological information. **Comment**: The definition of trade secret excludes trivial information and the experience and skills gained by employees in the normal course of their employment, and also excludes information which is generally known among, or is readily accessible to, persons within the circles that normally deal with the kind of information in question[41].
Confidential data	For BIPED purposes, intermediate versions of BIPED consortium project data and datasets are deemed confidential, irrespective of the license that the consortium establishes for final datasets.	It is important to distinguish this concept from the concept of "data confidentiality and integrity", trade secrets, and other intellectual property (IP) related concepts.
Aggregate data	Aggregation refers to a data mining process in statistics. Information is only viewable in groups and as part of a summary, not per the individual. Aggregate data may, but also may not be personal data depending on the circumstance[42].	**Comment:** Aggregate-level data is useful for answering research questions about populations or groups of people. This reduces privacy risks, but aggregation of a sample that is too small can lead to privacy issues. GDPR puts emphasis on the fact that aggregate data, statistical results or the personal data are not used in support of measures or decisions regarding any particular natural person.[43] **Example:** aggregate counts of people in an office space can be used in combination with other data, such as weather data, to create an energy-efficiency program so consumption is controlled, with the goal of saving money and reducing energy use.
Anonymized/de-identified data	anonymisation means the process of changing data/documents into anonymous data/documents which do not relate to an identified or identifiable natural person, or the process of rendering personal data anonymous in such a manner that the data subject is not or no longer identifiable[44].	**Comment:** To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the

[40] Article 2(1) of the Directive 2016/943 (EU) on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

[41] Ibid, recital 14.

[42] U.S. Federal Trade Commission: Is aggregate data always private? (Available at https://www.ftc.gov/news-events/blogs/techftc/2012/05/aggregate-data-always-private).

[43] Recital 162 GDPR.

[44] Article 2(7) of the Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

		available technology at the time of the processing and technological developments[45]. Sufficiently robustly anonymized data are not personal data.
Pseudonymized data	“Pseudonymisation” means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person[46].	**Comment:** Pseudonymization is not a method of anonymization. It merely reduces the linkability of a dataset with the original identity of a data subject, and is accordingly a useful security measure. Pseudonymized data is personal data. Encryption can be considered among the pseudonymization techniques.

Table 2: Data according to origin and purpose

Category	Definition
Original data	Data produced by a BIPED partner partner (e.g., data collected from sensors for BIPED purposes or data created during a dissemination action or a pilot activity).
BIPED existing data	Existing data already in possession of a BIPED partner prior to the project’s initiation. (Typically also Background information in the sense of Grant Agreement/Consortium Agreement).
Existing third party data	Data sourced/procured by a BIPED partner during the project’s timeline.
Soft data	Data in the form of qualitative information or quantitative information resulting from an approximation of economic phenomena through surveys and polls or other techniques (such as social media scraping).
IoT data	Smart cities use numerous resources such as sensors, cameras, mobile devices, etc. to collect data, route them through gateways and networks and eventually story them in a database.
Historical IoT data	Type of sensor data. The historical data set is a large volume of data typically indexed according to time and geographical dimensions. The historical data is mainly used to train digital twin models and visualize the past.
Context IoT data	The context data contains values as currently measured by the different devices. The context data is used as input for simulations and to visualize the present.
Location data	Data indicating the geographic position of a person or the terminal equipment of a person (user)[47].
Modeling data	Modeling data contains all data related to models and interactions.

[45] Recital 26 GDPR.

[46] Article 4(5) GDPR.

[47] Recital 14 of Directive 2002/58: Location data may refer to the latitude, longitude and altitude of the user's terminal equipment, to the direction of travel, to the level of accuracy of the location information, to the identification of the network cell in which the terminal equipment is located at a certain point in time and to the time the location information was recorded.

Smart data	Data or datasets extracted from larger amounts of data (big data, raw data) using algorithms according to certain structures, in order to provide meaningful information, understandable to the user, in order to help users achieve meaningful results. They may combine data coming from sensors, social media and other human-related sources and thus may contain personal data, including special categories of personal data.
Provided personal data	Data provided directly by the individuals concerned (such as responses to a questionnaire).
Observed personal data	Data observed about the individuals (such as location data collected via an application).
Derived/inferred personal data	Derived or inferred data such as a profile of the individual that has already been created (e.g. a credit score).
Dynamic data	Data/documents in a digital form, subject to frequent or real-time updates, in particular because of their volatility or rapid obsolescence; data generated by sensors are typically considered to be dynamic data[48].
Research data	Data/documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results[49].
Metadata	Metadata is data that provides information about other data[50]. For example, draft ePrivacy Regulation defines electronic communications metadata as "data processed by means of electronic communications services for the purposes of transmitting, distributing or exchanging electronic communications content; including data used to trace and identify the source and destination of a communication, data on the location of the device generated in the context of providing electronic communications services, and the date, time, duration and the type of communication"[51]. According to ISO, geospatial metadata "provides information about the identification, the extent, the quality, the spatial and temporal aspects, the content, the spatial reference, the portrayal, distribution, and other properties of digital geographic data and services"[52].

Table 3: Data processing operations

Data operation	Description

[48] Article 2(8) of the Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

[49] Article 2(9) of the Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

[50] https://en.wikipedia.org/wiki/Metadata .

[51] Proposal for a regulation concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications), European Council, 6 March 2020.

[52] ISO 19115:2013 "Geographic Information – Metadata".

Collection	Acquiring/creating data by asking questions and collecting responses (including via an online form), collecting data from sensors (other than recording), scraping the web.
Recording	Acquiring/creating over data by recording natural persons by means of audio-visual recording, taking photos, recording by a dictaphone, recording phone calls, keeping record of a meeting, recording that you have a person's consent for a particular type of processing of personal data.
Obtaining data from a third party data provider – open data/public information	Data accessible and open without any restrictions, or data accessible by an unlimited number of interested parties subject to applicable licensing conditions (open license access).
Obtaining data from a third party data provider/vendor (non-open data)	Purchasing data or otherwise individually negotiating access to data (individual access license).
Organisation and structuring	Sorting/grouping of data according to certain characteristics or logic, creating a filing system, creating a database.
Storage	Storing data in physical depositories or in the cloud, keeping data for longer, e.g. not erasing the data after they had been processed for a respective task. This can involve pseudonymization or encryption of data for secure storage.
Adaptation or alteration	Changing the nature, contents, quality of the data or metadata by correcting errors or updating the data. Typically done in order to maintain data accuracy. This activity may be done by you, but you may also allow users to alter data related to them via access to a personal account on a website.
Consultation and use	Use of data for making a decision, drawing a conclusion, forming an opinion, use of data (feeding) in algorithmic decision making or machine learning operations and systems.
Disclosure by transmission	Sharing of data with other organisations (other companies or authorities), but also within your organisation with different branches/sections/departments. This may include uploading of data to a cloud drive.
Dissemination or otherwise making available	Disclosure to the public or a wider group of recipients by means of e.g., webpage, generally available APIs, open database.
Alignment or combination	Integration or combination of data in a dataset (pre-existing or new), alignment of data so two datasets can interact.
Restriction	Marking of stored data with the aim of limiting their processing in the future[53].
Erasure or destruction	Erasure of data which are no longer needed for the envisaged purpose; removal from search index. Can be done individually or en masse, ad hoc or at regular points where different categories of data get erased. This can involve destruction of physical documents, media or hard drives.

[53] Article 4(3) GDPR.

Table 4: Other data management and privacy related concepts

Concept	Description
Data subject	Identified or identifiable natural person, subject of personal data related to that person.
Data owner	Legal entity or person that is ultimately responsible for the data governance regarding the data. See also Controller for a definition under the GDPR (applicable to personal data only).
Privacy by design	Implementation, at the time of the determination of the means for processing and at the time of the processing itself, of appropriate technical and organisational measures to protect the rights of data subjects[54].
Privacy by default	Implementation of appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility[55].
Filing system	A structured set of personal data which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis[56].
Personal data breach	Breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed[57].
Terminal equipment	Equipment directly or indirectly connected to the interface of a public telecommunications network to send, process or receive information; in either case (direct or indirect), the connection may be made by wire, optical fibre or electromagnetically; a connection is indirect if equipment is placed between the terminal and the interface of the network[58]. Examples: mobile phones, laptops, tablets, connected devices (connected vehicles).
Data Protection Officer/DPO	A person tasked by an organisation (a data controller) with ensuring compliance with the GDPR and other privacy related tasks.
GDPR	Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
Controller	The natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined

[54] Article 25(1) GDPR.

[55] Article 25(2) GDPR.

[56] Article 4(6) GDPR.

[57] Article 4(12) GDPR.

[58] Article 1(1) of the Directive 2008/63 on competition in the markets in telecommunications terminal equipment.

	by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law[59].
Processor	Natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller[60].
Data Protection Impact Assessment/DPIA	An assessment to evaluate the origin, nature, particularity and severity of risk to the rights and freedoms of natural persons[61].
Profiling	Any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements[62].
Machine-to-machine service/IoT services	Services involving an automated transfer of data and information between devices or software-based applications with limited or no human interaction[63]. These services may be considered an "electronic communication service"[64] and thus be subject to ePrivacy laws.
MIMs	Minimal Interoperability Mechanisms

Table 5: Dictionary / Glossary of terms used in BIPED Data Collection Sheet

Term	Definition
Scope	The extent of the area or subject matter that something deals with.
Theme	The relevant project theme
KPI Number	A unique identifier for a specific Key Performance Indicator (KPI).
Definition	A clear and precise description of the KPI.

[59] Article 4(7) GDPR.

[60] Article 4(8) GDPR.

[61] Recital 84 GDPR.

[62] Article 4(4) GDPR.

[63] Recital 12 of the proposal for a regulation concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications), European Council, 6 March 2020.

[64] Article 2(4) of the Directive (EU) 2018/1972 establishing the European Electronic Communications Code.

KPI Owner	The KPI owner takes the lead in the implementation, testing and monitoring of the project interventions. The KPI owners use the KPI framework created for the BIPED project to ensure that interventions are recorded and made available for analysis. The KPI owner will agree to the definition, description and calculation method of the KPIs, in cooperation with WP4. The KPI owner is responsible for implementing measures which will enable data to be captured, and providing this data in a suitable and agreed upon format, for example the M&E quantitative and qualitative data collection sheets, for reporting within the WP4 deliverables/updates and overall project reporting KPI owners are responsible for the completion of the data collection sheets according to the agreed upon reporting frequencies for each KPI and the partner responsible for the management and update of the SRT. Throughout the BIPED project the KPI owners will review the accuracy of data recorded and issue recommendations to the project consortium for adjusting the KPI definition and KPI calculations.
KPI technical experts	KPI technical experts are parties that act as complementary partners to KPI owners. KPI technical experts are specialists in their area/sector and provide technical support, tools and data to KPI owners which will assist in implementing project interventions. This support will contribute to the achievement of the KPI as well as providing trusted information which allows KPI owners to monitor and report on the data. KPI technical experts are responsible for the management of data from project interventions. KPI technical experts have the responsibility to handle data according to the Data Management Plan (DMP) and ensure that the handling of data adheres to best practice in data governance in accordance with protocols from Horizon Europe.
Data Type and Format	the data type (e.g. number, percentage etc) and how the data is presented in a certain format (e.g. CSV)
Data Source/Provenance	The origin of the data or where it was obtained. Differentiate between these categories: Existing data - third party provided / Existing data owned by BIPED partner / New data to be created/collected by a BIPED partner
Unit of Measurement	The standard unit in which the KPI is measured.
Associated Demonstration Project	A project/event that demonstrates the KPI in action.
Year of Data	The year in which the data was collected
Considerations	Factors or aspects that should be taken into account.
Expected Impact / Target	The anticipated impact of the KPI and the target to be reached.
KPI Share	The portion or percentage of the KPI that is shared across KPI owners.
Size	The magnitude or extent of the KPI or data.
Data Utility Outside BIPED	The usefulness or applicability of the data beyond the BIPED framework.
Quality and Validity	The degree to which the data is accurate, reliable, and valid.
Statistics Data	Data that has been collected for statistical analysis.
ISO Applied	Whether or not International Standards Organization (ISO) standards have been applied.

Lineage	The history or lifecycle of the data, including where it originated and how it has been altered over time.
Disclosure Control Methods (e.g. GDPR)	Methods used to control the disclosure of data, such as those outlined in the General Data Protection Regulation (GDPR).
Quality Issues	Any problems or issues related to the quality of the data.
KPI Owner / Organisation	The organisation that the KPI owner belongs to.
Organisation Name	The name of the organisation.
Email Address	The email address of the contact person in the organisation.
Responsible Party Role	The role of the person responsible for the data or KPI.
Telephone Number	The contact telephone number of the responsible party.
Resource Locator	The location where the resource can be found, often a URL.
KPI Owner Approval	Whether or not the KPI owner has approved the data or KPI.
Data Owner / Organisation	Legal entity or person (BIPED partner) that is ultimately responsible for the data governance regarding the dataset - if different from the KPI owner. In GDPR terms it is the data "Controller" for purposes of BIPED project.
Where Stored	The location where the data is stored.
Additional Solutions Providers	Any additional organisations providing solutions related to the data or KPI.
Temporal	Pertaining to time-related aspects of the data or KPI.
Temporal Extent	The time period that the data covers.
Frequency of Update	How often the data is updated.
Frequency of SCIS Update	How often the Smart Cities Information System (SCIS) is updated.
Dataset Reference Date	The date that the dataset refers to.
Planned Date of Implementation	The date when the implementation of the KPI or data usage is planned.
Actual Date of Implementation	The date when the implementation of the KPI or data usage actually occurred.
Monitoring Start Date	The date when monitoring of the KPI or data began.
Geographic	Pertaining to geographical aspects of the data or KPI.
Geography / Spatial Scale	The geographical area that the data covers.
Spatial Reference System	The coordinate system used to define geographical data.
Constraints	Any limitations or restrictions on the data or KPI.
Limitations on Public Access	Any restrictions on the public's access to the data.
Use Constraints	Any restrictions on how the data can be used.
Licence Type	The type of license that governs the use of the data.
Data provider Name	Name of the legal entity or person (third party, not a BIPED partner) that created / originated the dataset or is otherwise in possession of the data set and has rights to further provide or re-use the data under a licence.
Email Address	The email address of the contact person in the organisation.

Telephone Number	The contact telephone number of the responsible party.
Resource Locator	The location where the resource can be found, often a URL.
Where Stored	The location where the data is stored.
Conformity	Whether the data conforms to certain standards or expectations.
Metadata	Data that provides information about other data.
Metadata Date	The date when the metadata was created or last updated.
Metadata Language	The language in which the metadata is written.
Metadata Point of Contact	The person or organisation to contact for more information about the metadata.
Unique Resource Identifier	A unique identifier for the resource, often a URL.
Resource Type	The type of resource, such as a dataset, image, document, etc.
Dataset Language	The language in which the dataset is written.
Search Keywords	Keywords used to search for the data or resource.
Interoperability Best Practice	Best practices for ensuring that systems can work together (interoperate).
Vocabularies / Ontologies	Standardized vocabularies or ontologies used in the data. If the metadata are DCAT AP standard, please state so.
GDPR	Pertaining to the General Data Protection Regulation, a regulation in EU law on data protection and privacy.
Personal Data	Data that relates to an identifiable individual.
Special Categories of Personal Data	Categories of personal data that are considered sensitive under the GDPR or similar: personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.
Mixed Data	Data that includes a mix of personal and non-personal data.
Anonymisation/Pseudonymisation	The process of making data anonymous or pseudonymous to protect privacy.
Artificial Intelligence	The use of artificial intelligence in relation to the data or KPI.
AI Elements in the Model/Tool	Specific elements of artificial intelligence used in the model or tool.
Data Used to Train a Model	The data used to train a machine learning model.
Ethical	Pertaining to ethical considerations in relation to the data or KPI.
Ethical Considerations / Limitations	Any ethical considerations or limitations related to the data or KPI.
Envisaged Combination with Other Data/Sets/Models	Any plans to combine the data with other datasets or models.

Annex 2 - BIPED Data and Metadata Collection Sheet template

Scope	
Name	
Theme / Group	
KPI number	
Definition	
KPI owner	*BIPED Partner*
Data owner	*BIPED parnter that is ultimately responsible for the data governance regarding the dataset - if different from the KPI owner. In GDPR terms it is the "data controller".*
Data type and format	
Data source/provenance	*Existing data - third party provided / Existing data owned by BIPED partner / New data to be collected by a BIPED partner*
Unit of measurement	
Associated demonstration project	
Scope	*Number of All usual residents aged 16 to 74 (excluding students) in employment and currently working in Local Government District; Location of Usual Residence by Place of Work*
Year of Data	*2019*
Considerations	*1. The workplace population in an area does not include those persons working in the area who live outside Ireland.* *2. No fixed place is counted as if working in the area.*
Expected Impact / Target	
KPI Share	
Size	
Data utility outside BIPED	
Quality and Validity	
Statistics Data	*Yes*
ISO applied	
Lineage	*Census data are aggregated within different boundaries essentially by assembling small geographical building bricks to which the data are coded. These allow the data to be aggregated to a range of geographic units.*
Disclosure Control Methods (e.g. GDPR)	*The confidentiality of personal information is paramount, and disclosure protection measures are used to prevent the inadvertent disclosure of information about identifiable individuals. All outputs have been derived from a database within which the records have been subjected to statistical techniques to minimise the risk of inadvertent disclosure. In addition, broad limitations are placed on details in tables to be produced for small populations. There were minimum thresholds of numbers of person and households for the release of sets of output.*

Quality Issues	*The questionnaire including the questions asked and the administrative procedures involved in collecting the census data underwent substantial testing. Coding of the data was subject to quality checks. The quality of the results was improved by the use of edit and imputation procedures for missing or incorrect data, and the data were adjusted for under-enumeration. The results underwent an extensive quality assurance process, which included checks against administrative data sources and information on particular groups. Edit procedures were applied to obviously incorrect responses (such as someone aged 180) and were designed to correct the mistake by making the least possible change to the data. Imputation procedures were applied to missing data on a returned form, and drew on responses to the question from people with similar characteristics.*
KPI Owner (BIPED Partner)	
Organisation Name	
Email Address	
Responsible Party Role	
Telephone Number	
Resource Locator	*www.data.gov.ie*
Where stored	*trusted repository*
KPI Owner Approval	
Data Owner (BIPED partner if different from KPI Owner)	
Organisation Name	
Email Address	
Responsible Party Role	
Telephone Number	
Resource Locator	*www.data.gov.ie*
Where stored	*trusted repository*
Additional Solutions Providers	
Organisation Name	
Email Address	
Responsible Party Role	
Telephone Number	
Resource Locator	
Temporal	
Temporal Extent	*Day: 27 January 2019*
Frequency of Update	*Once a year*
Frequency of SCIS Update	*Once a year*
Dataset Reference	*2019*

Date	
Planned date of implementation	
Actual date of implementation	
Monitoring start date	
Geographic	
Geography / Spatial Scale	*SCD, City etc*
Spatial Reference System	*Irish National Grid*
Data provider and Constraints	
Limitations on Public Access	*No restriction on public access*
Use Constraints	*No conditions apply*
Licence type	
Data provider Name	
Email Address	
Telephone Number	
Resource Locator	*www.data.gov.ie*
Where stored	*trusted repository*
Conformity	
Conformity	*See Statistics Status*
Metadata	
Metadata Date	*2018-06*
Metadata Language	*eng*
Metadata Point of Contact	
Unique Resource Identifier	*code*
Resource type	*Dataset*
Dataset Language	*eng*
Search keywords	
Interoperability best practice	
Vocabularies / ontologies	*Note whether DCAT- AP standard or other vocabulary*
GDPR	
Personal data	*Yes/No*

Special categories of personal data	*(personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person’s sex life or sexual orientation)*
Mixed data	*(personal and non-personal data in one dataset)*
Anonymisation/pseudonymisation	
Artificial Inteligence	
AI elements in the model/tool	*yes/no, describe*
Data used to train a model	*yes/no, describe*
Ethical	
Ethical considerations / limitations	
Envisaged combination with other data/sets/models	

Annex 3 - Model Information Sheet and Consent Form (Privacy)

Model BIPED Study Information Sheet

BIPED Project Study Information Sheet

INTRODUCTION
You are invited to join a research study in the BIPED project. Before you decide to participate in this research study, it is important that you understand why the research is being done and what it will involve. Please read the following information carefully and take whatever time you need to discuss the study with your colleagues and friends, or anyone else you wish to. The decision to join is on a voluntary basis.
Please ask the researchers if there is anything that is not clear or if you need more information.

BIPED RESEARCHERS
BIPED project is implemented by the following entities (hereinafter referred as "**BIPED Team**") in accordance with the terms set out in the Grant Agreement signed with the Research Executive Agency (REA):

1. [list BIPED partners]

PURPOSE OF THE RESEARCH STUDY
BIPED is funded under the EU Horizon Europe Research and Innovation programme. Grant ID: 101139060.

Positive Energy Districts (**PEDs**) are a key building block in the future energy paradigm for carbon-neutral cities and communities. In this research study, we are investigating how local digital twins (**LTDs**) can be extended to refine a district's profile representation, how soft data can support the advancement of digital twin development and how to boost PED's replication potential for the benefit of supporting smart cities.

[*BIPED partners: please insert a specific but clear description of the purpose of the survey in the context of the BIPED research activities.*]

WHAT IS INVOLVED IN THE STUDY?
If you decide to participate you will be asked to participate in a survey questionnaire *[please check and add other info if needed]*.... We think this will take you_____ minutes.

PERSONAL DATA
By participating in the study,
You understand that the personal data that will be gathered for the purposes of this survey is: [Name and surname; Name of your organization; E-mail address].

You consent to its collection, processing, and storage by the BIPED Team on the basis of information provided under this privacy statement, and in compliance with the limitations set out below.

You understand that the data that identifies you and that you consent to provide by participating in this survey, can be shared with the third parties only if expressly mentioned and identified in the Informed consent form, if any, and that this data will be treated in full compliance with the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (**GDPR**) by the applicable party. No other third party will have access to your personal data.

BENEFITS

The study may help us understand how we can best use LTDs and data to help positive energy districts, the municipalities and their citizens. We will share the results of our survey with you in the form of the project deliverables.

RISKS

Due to the minimal risk of this research we will not ask you to sign a formal contract but participating in the research implies acceptance of the conditions stated above.

CONFIDENTIALITY

We will take the following steps to keep information about you confidential, and to protect it from unauthorized disclosure, tampering, or damage:

- Your opinion or data will be anonymized, or used only as aggregated statistical data.
- All data is stored on a secured server environment and is only accessible to the BIPED researchers.
- Publications will not mention your name or function unless specifically requested and agreed by you.

YOUR RIGHTS AS A RESEARCH PARTICIPANT

Participation in this study is voluntary.

You have the right not to participate at all or to terminate your involvement at any time. Deciding not to participate or choosing to leave the study will not harm your relationship with the partners of BIPED consortium.

If you decide to take part in this study, you will be asked to sign a consent form. After you sign the consent form, you are still free to withdraw at any time and without giving a reason.

PUBLICATION OF RESEARCH FINDINGS

You can follow the progress and results of our project on the website https://www.bi-ped.eu/.

CONTACTS FOR QUESTIONS OR PROBLEMS

Please call [..] or email [..] if you have questions about the study, any problems, or think that something unusual or unexpected is happening.

We thank you a lot for participating in this project!

The BIPED Team

Model BIPED Informed Consent Form

Informed Consent for BIPED project survey

Participating in the BIPED project research

a) I have read and understood [*BIPED Study Information Shee*t] dated […./…./…..], or it has been read to me. I have been able to ask questions about the survey and my questions have been answered to my satisfaction.

b) I consent voluntarily to be a participant in this survey, further explained in [*BIPED Study Information Shee*t] and understand that I can refuse to answer questions and I can withdraw from the research at any time, without having to give a reason.

Personal data in the context of the current survey

c) I explicitly consent to this data □ name and surname, □ e-mail, □ organisation, about me being collected, treated and processed, the context of the survey by *[x, y, z. BIPED partners: please complete who will be the organisation which collects this data in the context of the survey]*.

d) I consent that this data □ name and surname, □ e-mail, □ organisation, will be used only for the purpose of *[□ x, □ y, please complete]*.

Future use and reuse of the personal information by third parties

e) I understand that the data □ name and surname, □ e-mail, □ organisation, that identifies me and that I consent to provide by participating in this survey, can be shared with *[BIPED partners: please list all BIPED partners that will receive non fully anonymised data]* and that this data will be treated in full compliance with the GDPR by those parties.

Data retention and destruction policy

f) I consent that this data □ name and surname, □ e-mail, □ organisation, can be retained by [*DATA PROCESSOR to be added*] in the survey database for *[x time, please complete*[65]], and collected and processed on the basis of the above limitations, after which it will be destroyed.

I have read this informed consent form and agree to participate in the survey.

□ YES □ NO

[65] provide this information, taking into account that according to Art. 4 of the GDPR personal data shall be kept for no longer than is necessary for the purposes for which the personal data are processed and may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by the GDPR in order to safeguard the rights and freedoms of the data subject ('storage limitation').

Signature

______________________________ ________________________

Name of the participant Signature Date

BIPED project contact details for further information

https://www.bi-ped.eu/
hello@bi-ped.eu