
# “Building the Next Generation Personal Data Platforms”
# G.A. n. 871370

**Deliverable D7.6**

DMP and report on RRI and societal impact

**H2020-EU-2.1.1: PIMCity**

**Project No. 871370**

**Start date of project: 01-12-2019**

**Duration: 33 months**

**Revision:** 01

**Delivery Date:** 31-08-2022

**Document Information**

**Document Name: Data Management Plan and report on RRI and societal impact**

**WP7 – Title: Data management and legal and ethical requirements**

**Tasks 7.1, 7.3**

**Authors:** Enzo Marquet (KUL) as main editor, and all partners involved in WP7.

**Dissemination Level**

| **Project co-funded by the EC within the H2020 Programme** | | |
|---|---|---|
| **PU** | Public | ☑ |
| **PP** | Restricted to other programme participants (including the Commission Services) | |
| **RE** | Restricted to a group specified by the consortium (including the Commission Services) | |
| **CO** | Confidential, only for members of the consortium (including the Commission Services) | |

(Tick the corresponding dissemination level of the deliverable according to Annex I).

**Approvals**

| | **Name** | **Entity** | **Date** | **Visa** |
|---|---|---|---|---|
| **Authors** | 1. POLITECNICO DI TORINO, the Coordinator (Marco Mellia – Martino Trevisan – Tiziana Rolando – Giorgio Prette)<br>2. NEC LABORATORIES EUROPE GMBH (Roberto Gonzales - David Friede)<br>3. ERMES CYBER SECURITY SRL (Stefano Traverso)<br>4. FUNDACION IMDEA NETWORKS (Alvaro Garcia, Nikolaos Laoutaris, Santiago Andrés Azcoitia, Devris Isler)<br>5. UNIVERSIDAD CARLOS III DE MADRID (Amir Mehrjoo, Angel Cuevas, Antonio Pastor, Rubén Cuevas)<br>6. TELEFONICA INVESTIGACION Y DESARROLLO SA (Ioannis Arapakis, Panagiotis Papadopoulos<br>7. FASTWEB SPA (Guglielmo Bondioni)<br>8. LSTECH ESPANA SL (Evangelos Kotsifakos, Xavi Olivares) | **All partners** | **31-08-2022** | |

| | | | | |
|---|---|---|---|---|
| | 9. KATHOELIEKE UNIVERSITEIT LEUVEN (Enzo Marquet, Viltė Kristina Dessers, Peggy Valcke)<br>10. ASOCIACION DE USUARIOS DE INTERNET (Miguel Pérez Subías)<br>11. INTERACTIVE ADVERTISING BUREAU SPAIN (Miguel Herranz – Paula Ortis)<br>12. WIBSON (Rodrigo Irarrazaval – Daniel Fernandez)<br>13. TAPTAP (Álvaro Mayol) | | | |
| **WP Leader** | **Enzo Marquet, Viltė Kristina Dessers, Peggy Valcke** | **KUL** | **31-08-2022** | |
| **Coordinator** | **Marco Mellia** | **POLITO** | **31-08-2022** | |

**Document history**

| Revision | Date | Modification |
|---|---|---|
| **Version 1** | 30-06-2022 | Initial version |
| **Version 2** | 29-08-2022 | Revised and completed draft |
| **Version 3** | 31-08-2022 | Final version |

**LIST OF ABBREVIATIONS**

| Abbreviation | Meaning |
|---|---|
| **D7.1** | Deliverable D7.1 of the PIMCity project |
| **D7.2** | Deliverable D7.2 of the PIMCity project |
| **D7.3** | Deliverable D7.3 of the PIMCity project |
| **D7.4** | Deliverable D7.4 of the PIMCity project |
| **D7.5** | Deliverable D7.5 of the PIMCity project |
| **DMP** | Data Management Plan |
| **DPIA** | Data protection impact assessment |
| **DPO** | Data protection officer |
| **EU** | European Union |
| **FAIR data** | Findable, accessible, interoperable and re-usable data |
| **GDPR** | General Data Protection Regulation |
| **PIMCity project** | Horizon2020 project PIMCity (Building the next generation Personal Data Platforms), Grant Agreement No. 871370 |

**TABLE OF CONTENTS**

## INTRODUCTION

The updated PIMCity Data Management Plan (further as the DMP) sets out the final data management plan for the PIMCity project and reflects the consortium's comprehensive approach and joint efforts towards making research data findable, accessible, interoperable and re-usable (further as FAIR). The DMP describes the data management life cycle for the data to be collected, processed and/or generated by the PIMCity project. As part of making research data FAIR, the DMP provides the information on the handling of research data during and after the end of the project. The DMP indicates what data are collected, processed and/or generated, which methodology & standards are applied, whether data are shared/made open access and how data is curated & preserved (including after the end of the project)[1]. Content of the DMP is consists of the information provided by the project partners as indicated in the relevant sections. The DMP is drafted in accordance with the Guidelines on FAIR Data Management in Horizon 2020 as of 26th July, 2016, issued by the European Commission Directorate-General for Research & Innovation.

The document also includes the updated and final version of the data protection impact assessment (further as the DPIA) as well as various guidelines and template forms that shall be adjusted by the partners on a case-by-case basis taking into account particular circumstances.

## STRUCTURE

The DMP consists of two parts. The first part clarifies the general approach towards data management as adopted by the project partners. The second part consists of partners' individual inputs which clarify the approach of each partner individually (Annex H). The DPIA, guidelines and templates are added as Annexes A-G.

## DISCLAIMERS

The information provided in the DMP, guidelines and templates (Annexes A-H) does not constitute legal advice. Any user of this information uses it at its sole risk and liability.

## PIMCITY DATA MANAGEMENT PLAN

The PIMCity project collects and/or generates at least four broad categories of data as provided in the table below. The information provided further in the DMP clarifies this information further. Particularly, the DMP elaborates on the management of these broad categories of data and reveals in detail the relation of each

[1] European Commission Directorate-General for Research & Innovation. Guidelines on FAIR Data Management in Horizon 2020 as of 26th July, 2016, p. 4.

dataset with the PIMCity project's exploitable outputs which allows gaining a clear understanding of the limitations of the OA as well as the impact on the exploitation and dissemination strategies.

| Category | Format | Means for OA | Curation and cost allocation |
|---|---|---|---|
| **Documents and dissemination materials**: Includes deliverables, reports, demonstrations, dissemination and communication | Common text, image or video formats (.pdf, .docx, .jpeg, .mov, .avi, etc.) | Self-archive on website; green scientific publications + OPENAIRE repositories | Technical coordinator (T8.2); dissemination manager (T6.1); peer review: scientific journal panels |
| **Computer software**: including software applications (in binary form), libraries in the form of SDKs, plugins and respective source code | Binary format, ZIP archives; Source code in common files such as C, CPP, etc. | GitHub | Technical coordinator (T8.2); innovation manager: exploitability and license schemes (T6.4) |
| **Research data and metadata**: materials and datasets resulting from the implementation of the developments; metadata and configuration files; bug logs and feedback logs; developer internal documentation; evaluation and opinions | Log files; text files using (.pdf, .docx, .xls, etc.) | green scientific publications + OPENAIRE repositories | Innovation manager: exploitability (T6.4); data manager: anonymization of evaluation questionnaires and opinions; conditions of pre-existent data (T7.1); dissemination manager (T6.1) |
| **Data for evaluation**: consists in materials or datasets generated or collected by the project used for evaluation purposes | Log files; files using (.docx, .xls, etc.) | Green scientific publications + OPENAIRE repositories | Data manager: conditions of pre-existent data (T7.1); dissemination manager (T6.1) |

**Approach to personal data**

Some of the project partners had, have and/or will have access to personal data. All the data were, are and will be processed in accordance with the relevant EU legal requirements. Detailed information about personal data processing within the PIMCity project and the platform created in the framework of this project is available in D7.5 as well as in the DPIA provided further in this deliverable.

The PIMCity partners appointed a data protection officer team to act as a single point of contact for data subjects wishing to exercise their rights, following the provisions of the Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data[2] (further as the GDPR) and taking into account the Article 29 Working Party Guidelines on Data Protection Officers.[3]

Partners also implemented all the necessary technical means to ensure data subject's rights such as the right to withdraw consent, right of access, right to erasure, right to data portability, including and not limited. More details on various technical measures are provided in D9.2 (confidential).

[2] Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.
[3] Article 29 Working Party Guidelines on Data Protection Officers, adopted on 13 December 2016, last revised and adopted on 5 April 2017 https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612048, and endorsed by the European Data Protection Board which replaced the Article 29 Working Party on 25 May 2018.

The PIMCity partners confirm that eventually no personal data were processed outside the EU.

**Re-use and open access**

Overall, the PIMCity project partners generated a significant amount of research data with potential for re-use and verification. As identified in the PIMCity project proposal, in compliance with Responsible Research and Innovation on Open Access (further as OA), default policy is to make its data publicly available through public copyright licenses (e.g. Creative Commons, etc.). Possible solution can be, among others, archiving it on the project website and OPENAIRE compliant repositories. However, with regard to certain data PIMCity partners may need to apply OA restrictions. The latter may stem from, including but not limited, (i) confidentiality and intellectual property protection of certain deliverables, datasets and outputs (e.g. underlying algorithmic and methods susceptible of being patented); (ii) protection of personal data of persons involved in feedback collections; (iii) protection of imported rights of pre-existent, non-public datasets.

| **1. DATA SUMMARY** | *Details on (i) the purpose of the data collection/generation and its relation to the objectives of the project; (ii) the types and formats of data the project generates/collects; (iii) the plans to re-use any existing data and, if yes, details on how; (iv) the origin of the data; the expected size of the data; (v) to whom might it be useful ('data utility').* |
|---|---|
| | (i) In general, data is collected and generated for the purposes of the project which are, including but not limited, to create privacy preserving tools. In particular, the project partners aim to implement a Personal Information Management Systems (PIMS) development kit (PDK) to commoditize the complexity of creating PIMS. Second, project partners aim to design and deploy novel mechanisms to increase users' awareness: (i) the Transparency Tags (TT) which would show users essential information about the services they access, in a simple and easy to understand manner; (ii) the Personal Data Avatar (PDA) which would be intuitive means for users to control the information shared to third parties. Third, the project partners aim to demonstrate the effectiveness of the above tools by engineering EasyPIMS, the fully-fledged PIMS. Overall, the project partners partners aim to build the largest-ever transparent data marketplace implementing and demonstrating EasyPIMS with a number of end-users, collaborating with advertisers and operators in the web market.<br><br>(ii) Overall, the partners intend to generate/collect these data:<br>- information instrumental to check if and how the website or web service collects and exchanges eventual personal information; privacy tags summarize the output of the algorithms (Politecnico di Torino);<br>- data composed by sequences of host and stored in a mysql database (NEC Laboratories Europe);<br>- data generated by its fleet of automatic web scrapers, which are then processed by automatic algorithms to identify which personal data web services collect and how; the result of this analysis, the D-PM, is provided in JSON format, and stored and distributed using state-of-the-art database technologies (Ermes Cyber Security);<br>- existing public data available on the Internet that may eventually resemble the kind of information that the PDK or the EasyPIMS would be managing and trading; this is in general structured data, for instance anonymized mobility data or CDRs (call detail records) showing the mobility of people within an area, sich as a city; more details could be provided once the specific use cases to implement are clear; as of now, we are working with this |

| | |
|---|---|
| | mobility information which in all cases is totally anonymized (IMDEA Networks; LSTech ESPANA);<br>- data regarding the value/price of audiences (i.e., user profiles) in online advertising platforms (Universidad Carlos III de Madrid);<br>- in most cases textual type with formats such as JavaScript Object Notation (JSON) and Comma Separated Values (CSV); depends on the format that the 3rd party systems provide their data (Telefónica Investigación y Desarrollo);<br>- aggregate network traffic data (e.g. traffic share of a particular website or service, bandwidth usage over time, number of users over time); the aggregation occurs in the dimension of users, i.e.: we do not collect or share any individual user identifier, but only user counts or session counts or bandwidth per website/service, possibly over time (Fastweb);<br>- identifying data (name, email, phone) collected exclusively to contact participants who wish to give them voluntarily; as for the formats these are collected in records and tables of mysql type databases in which the identification data are always stored encrypted (Asociación de Usuarios de Internet);<br>- sociodemographic data (age, gender, country, marital status, professional status, level of education, country, state, city, language) for the elaboration of studies and evaluation of the use of the different tools; as for the formats these are collected in records and tables of mysql type databases in which the identification data are always stored encrypted (Asociación de Usuarios de Internet);<br>- use and activity data (log files); as for the formats these are collected in records and tables of mysql type databases in which the identification data are always stored encrypted (Asociación de Usuarios de Internet);<br>- information from Data Sellers and Data Buyers as JSON, responding to the transaction history in JSON format; the raw input data format for the TG varies depending on the data source, e.g. the historical geolocation data can be a KML or a JSON file; the schema generated for the data source is served as JSON;<br>- on the scope of the consent manager (P-CM), the data consisst on tuples <identifier_of_user_data, consent_level> for each user; the consent_level is a structure common across all users; the identider_of_user_data is a pointer to arbitraty user's data stored on the P-DS; on the scope of the Trading Engine (TE), the data consists on contract, queries expresed by buyers; in the case of fullfillment of the contract, it is stored for auditing purposed; contracts contain no personal data but pointers to all relevant actors of it, namely: identider_of_user_data, identifier_of_seller, intefider_of_buyer and other metadata relevant for the contract such as date, experation, price, etc.;<br>- KU Leuven – CiTiP generates research data in the form of deliverables which is saved primarily in .docx and .pdf formats.<br><br>(iii) the project partners intend to re-use existing data, however most of these data shall not be personal data; in detail:<br>- Politecnico di Torino has historical web crawling archives that have been performed in the past; these contains a snapshot of web pages done through the time using automatic web crawlers; Polito acts as DP and is the only partner that can access the EasyPIMS data processor and data controller. It will keep the access for the next 5 years as required by the grant agreement.<br>- NEC Laboratories Europe re-uses anonimyzed dataset containing traffic logs; |

| | |
|---|---|
| | - Ermes Cyber Security re-uses data collected in the past using its fleet of automatic web scrapers. This contains information about web sites including the code used to generate the page (e.g., HTML and Javascript) as well as logs describing the APIs executed by the browser to generate the page;<br>- IMDEA Networks conducts the tests by reusing public data, by downloading copies of the target datasets to local development environments and using them in the testbeds;<br>- Universidad Carlos III de Madrid has some data collected before the start of the project, in the context of TYPES and SMOOTH EU projects and they might reuse it. It is of the same nature and type than the one they plan to collect in PIMCITY;<br>- Fastweb might re-use some of the aggregated network traffic data it has been collecting for network capacity planning purposes, if useful for processing in PIMCity modules;<br>- LSTech ESPANA uses data collected and used by IMDEA since both are participating in the same task. All the tests are done by reusing public data, by downloading copies of the target datasets to local development environments and using them in the testbeds;<br>- WIBSON takes into account that TE stores transactional Data for the duration of the action. During this period, data remains untouched, and both Data Sellers and Data Buyers can access it to see the history whenever they need to. The TG works similarly, receives raw input data, generates a schema, and serves it to the systems that need that information;<br>- Telefónica Investigación y Desarrollo, Asociación de Usuarios de Internet do not plan to re-use data;<br>- KU Leuven – CiTiP re-uses the generated research data.<br><br>(iv) in detail:<br>- Politecnico di Torino has some data collected by them internally by automatic web crawlers running on clusters of computers. Other are from public repositories such as https://web.archive.org. Politecnico di Torino plans to continue and enrich the web-crawling archives. The size of the web archives can be very large, up to several terabytes of data depending on the extensiveness of the collection, and the frequencies at which these are collected. The size of the privacy tags archive are much smaller, in the order of few gigabytes; Considering EasyPIMS, Politecnico di Torino is the only partner that has access to the data collected by the platform. It will continue handling such data for the next 5 years as by the grant agreement;<br>- NEC Laboratories Europe, Asociación de Usuarios de Internet deals with data provided by users; the size of the data to be clarified;<br>- Ermes Cyber Security has collected data internally, using its fleet of automatic web crawlers running on clusters of computers deployed in the cloud; the size of the data is not know yet, i.e. Ermes Cyber Security has collected a few terabytes of data so far, but such size varies depending on a number of variables (e.g, the amount of browsed sites, number of samples, iterations, etc.). The size of the resulting dataset containing D-PMs is expected to be much smaller (few GBs);<br>- Universidad Carlos III de Madrid handles data from advertising Platforms (closed ones as well as OpenRTB based); the size of the data shall be between ten and thousands of GBs of data;<br>- Telefónica Investigación y Desarrollo handles data of third party Personal Information Management (PIM) systems (e.g. Facebook, Mobile Phone, Email etc.); the size of the data shall be hundreds of megabytes per user; |

| | |
|---|---|
| | - the data processed in PIMCity modules hosted in Fastweb's cloud computing infrastructure comes from the sources of said modules, developed and controlled by other Consortium parties. The data collected directly by Fastweb comes from network sensors installed in Fastweb's customer network; the size of the data shall be determined later;<br>- IMDEA Networks and LSTech ESPANA public data available in search engines (e.g. Google datasets) or provided by public entities (e.g. https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page). Such datasets are made public without including identifiable personal data for developers and data scientist to test their algorithms; the size of the data strongly depends on the use case, ranging from MB to tens of GB (e.g. mobility data);<br>- WIBSON: the TE stores data from the Offers created by Data Buyers and the transactions that happen when Data Sellers accept selling their data; the TG receives raw input data from the Data Portability and Control tool; with regard to the size of the data, TE stores Offer and transactional data, and expect to generate less than 10GB; the schema generated by the TG occupies less than 1GB;<br>- KU Leuven – CiTiP generates research data relying on their own expertise in the subject matter.<br><br>(v) see (i). |

| **2. FAIR DATA** | |
|---|---|
| 2.1. Making data findable, including provisions for metadata | *Details on (i) whether the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism;[4] (ii) what naming conventions will the partners follow; (iii) whether search keywords will be provided that optimize possibilities for re-use; (iv) whether clear version numbers will be provided; (v) what metadata will be created; in case metadata standards do not exist, details on what type of metadata will be created and how.* |
| | (i) In general, indexes and standards are not relevant in the light of this project; data generally shall not be re-useable; datasets are used for specific goals, however, are not supposed to be useful for general use;<br>(ii) The partner agreed to use custom naming format given neither standard nor best-practice solutions are available for the type of data collected during the project;<br>(iii) see (i);<br>(iv) public documents have version numbers;<br>(v) The metadata includes the time of collection, the definition of the scheme of the data itself, the entity that collected the data, and the mechanisms used to collect the data. |
| 2.2. Making data openly accessible | *Details on (i) which data[5] will be made openly available as the default; (ii) whether certain datasets cannot be shared;[6] (iii) how will the data be made accessible;[7] (iv) what methods or software tools are needed to access the data; (v) whether documentation about the software is needed to access the data included; (vi) whether it possible to include the relevant software (e.g. in open source code); (vii) where will the data and associated metadata, documentation and code be deposited;[8] (viii) whether the partners have explored appropriate arrangements with the identified repository; (ix) whether there are restrictions on use, how will access be provided; (x) whether there is a need for a data access committee; (xi) whether there are well-described conditions for access;[9] (xii) how will the identity of the person accessing the data be ascertained.* |

[4] E.g. persistent and unique identifiers such as Digital Object Identifiers.
[5] Produced and/or used in the project.
[6] Or need to be shared under restrictions. If yes, explanation separating legal and contractual reasons from voluntary restrictions shall be included.
[7] E.g. by deposition in a repository.
[8] Preference should be given to certified repositories which support open access where possible.
[9] I.e. a machine-readable license.

| | |
|---|---|
| | (i) The consortium agreed to limit the open data to the minimum to avoid making public personal data of end-users.<br>(ii) In fact, all data collected by the EasyPIMS platform and all data used for the Telco demonstrator cannot be made open to avoid making public the personal data of participants;<br>(iii) in the project website(s) and on the Zenodo platform;<br>(iv) no specific access tools are necessary to access deliverables that are uploaded to the project website(s); software via standard repositories;<br>(v) yes; are provided;<br>(vi) project partners have committed to make all software open source; All the PDK, libraries, source code of demonstrators are accessible via the project website(s) and offered on the GITLAB repository;<br>(vii) the project partners use only internal servers and do not use any public repository or commercial cloud servers; Microsoft Teams is used for internal storing; the project website(s) – for external access;<br>(viii) The partners agreed to use the Fastweb Cloud service as the main platform for development and testing of their demonstrator;<br>(ix) everything shall be available for free with no restrictions except certain software (to be identified) and certain parts of the documents that may contain intellectual property, trade secrets or other information access to which shall be limited;<br>(x) the consortium agreed that a data access committee is not required;<br>(xi) see (ix); relevant documents are provided for accessing software;<br>(xii) people is be able to access most of the deliverables without revealing their identities and personal data of those people is not stored; for the purposes of statistics (downloads of the deliverables) only aggregated data shall be used (e.g. that certain amount of downloads were in Germany); certain software users are identified via licencing agreements. |

| | |
|---|---|
| 2.3. Making data interoperable | *Details on (i) whether the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc.;*[10] *(ii) what data and metadata vocabularies, standards or methodologies will the partners follow to make the data interoperable; (iii) whether the partners will be using standard vocabularies for all data types present in the data set, to allow inter-disciplinary interoperability; (iv) in case it would be unavoidable that the partners use uncommon or generate project-specific ontologies or vocabularies, will they provide mappings to more commonly used ontologies.* |
| | (i) There are no relevant standards and the partners did not engage into activities that focuses on creating new standards; various deliverables were indexed automatically;<br>(ii) The partner agreed to use custom naming format given neither standard nor best-practice solutions are available for the type of data collected during the project<br>(iii) no, see (ii);<br>(iv) no, see (ii). |
| 2.4. Increase data re-use (through clarifying licences) | *Details on (i) whether the data will be licensed to permit the widest re-use possible; (ii) when will the data be made available for re-use;*[11] *(iii) whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project;*[12] *(iv) how long is it intended that the data remains re-usable; (v) quality assurance processes.* |

[10] I.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins.

[11] If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

[12] If the re-use of some data is restricted, it shall be explained why.

| | |
|---|---|
| | (i) Data generally shall not be re-useable; datasets are used for specific goals, however, are not supposed to be useful for general use; only certain software shall be licenced;<br>(ii) within two months after generating particular data;<br>(iii) yes;<br>(iv) five years after the project ends; the project website(s) shall be available for five years after the project ends;<br>(v) N/A. |
| **3. ALLOCATION OF RESOURCES** | *Details on (i) what are the costs for making data FAIR in the project; (ii) how will these be covered;[13] (iii) who will be responsible for data management in the project; (iv) are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long).* |
| | (i) Costs are mostly represented by two main contributions:<br>- costs to manage and make data accessible for a Fair approach and<br>- costs for hosting data API in servers that hosts data and offer API to access it.<br>(ii) each partner is responsible for the costs incurred in making data FAIR, and for open access;<br>(iii) responsible partners are indicated in the table at the beginning of this document. Particular roles are foreseen for technical coordinator (T8.2); dissemination manager (T6.1); innovation manager (T6.4); data manager (T7.1). Besides, PIMCity data protection officer team, subject to their competences, advises the PIMCity project partners on data management;<br>(iv) The consortium selected the ZENODO platform which offer permanent and free data storage capability. Thus no cost is involved. |
| **4. DATA SECURITY** | *Details on (i) what provisions are in place for data security;[14] (ii) whether the data safely stored in certified repositories for long term preservation and curation.* |
| | (i) Each partner adopted individual data security measures;<br>(ii) Repositories were not used. |
| **5. ETHICAL ASPECTS** | *Details on (i) whether any ethical or legal issues can have an impact on data sharing;[15] (ii) whether the informed consent for data sharing and long-term preservation is included in questionnaires dealing with personal data.* |
| | (i) Defined in the deliverables of WP7 and WP9.<br>(ii) the project partners comply with all the relevant personal data protection requirements. More details are provided in D7.5. |
| **6. OTHER** | *Details on whether there is any use of other national/funder/sectorial/departmental procedures for data management and, If yes, references to the particular ones.* |
| | The project partners are following best industry practices and operate in compliance with internal procedures. |

[13] Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).
[14] Including data recovery as well as secure storage and transfer of sensitive data.
[15] These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

# 7. UPDATE ON THE REPORT ON RESPONSIBLE RESEARCH AND INNOVATION AND SOCIETAL IMPACT

**INTRODUCTION**

This part of the document includes an updated version of PIMCity Report on Responsible Research and Innovation[16] (further as the Report on RRI) as initially submitted as D7.3 on 22 March 2021[17] and updated as D7.4 on 1 March 2022.

The project partners had a rather clear idea about their approach to Responsible Research and Innovation (further as the RRI) since the beginning of the project, hence this final version of the Report on RRI entails only very minor changes compared to the approach already outlined in D7.3 and D7.4. This deliverable reconfirms the partners' approach to ensuring the RRI via particular elements of public engagement, Open Access, gender, ethics, science education.

**STRUCTURE**

The part of this deliverable which concerns the Report on RRI (part 4) is divided in the key sections of (1) public engagement, (2) Open Access, (3) gender, (4) ethics, (5) science education and (6) societal impact. Sections 1-5 provide both the general approach as adopted by the project partners and individual inputs of the partners which clarify the approach adopted by each partner. Section 6 provides the general approach as adopted by the project partners. The results of the Visual Law Lab Workshop (further as the Workshop) organised in M12 in order to improve accessibility, readability, accuracy and consistency of the privacy policy and informed consent forms that were drafted under D7.1. are provided in the public engagement section (1).

## 7.1. REPORT ON RESPONSIBLE RESEARCH AND INNOVATION

### 7.1.1. PUBLIC ENGAGEMENT

[16] RRI implies that societal actors such as researchers, citizens and businesses, including but not limited, shall work together during the research and innovation process in order to better align both the process and its outcomes with the values, needs and expectations of society. For example, in the light of PIMCity project, it shall provide key means to preserve the rights to privacy and data protection, ensure inclusive and socially acceptable Research. Overall, RRI shall contribute to providing durable societal impact. Responsible research & innovation. European Commission: online access at https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation [accessed on 2020-11-11].

[17] The Report on RRI reflects the consortium's comprehensive approach and joint efforts towards responsible Research and innovation as an approach adopted in the European Commission Horizon 2020 projects (further as RRI).

PIMCity plan for public engagement spans through the dissemination and communication activities, training and community building and the involvement in the co-design of the solution (primarily WP6 and WP1).

The partner *Asociación de Usuarios de Internet* (further as AUI) coordinates the activities (T6.1-6.4) that contributes to public engagement. A detailed strategy which provides comprehensive plan including dissemination, communication and community building strategies, as well as dissemination and exploitation reports are provided as deliverables of WP6 on the project's website[18] (see, e.g. D6.2, D6.3 and D6.4 as completed on May 2020, December 2020, and November 2021 accordingly, and D6.5, D6.6 and D6.7 completed at the end of the project).

Overall, PIMCity partners aim to contribute to building to a scientifically literate society whose members are well aware of their rights to privacy and data protection, of the developing data markets and the related legal and technological solutions. Furthermore, PIMCity partners aim for their Research to be future-proof and, in relation to this, for it to be supported by the particular knowledge and evidence of the societal challenges. In order to achieve these goals, PIMCity partners aim to engage the society into the development of the PIMCity solutions by establishing inclusive, participatory, and multi-actor dialogues between different stakeholders such as researchers, policymakers, industry and civil society organisations, NGOs, and citizens (e.g. via workshops as organised by KU Leuven – CiTiP, discussed in detail below).

PIMCity partners assume that it is important to ensure the public engagement from an early stage,[19] and they have already undertaken several actions during the first year of the project as revealed in the deliverables mentioned above and discussed briefly below.

*Input of different partners*

**KU Leuven – CiTiP** has contributed to public engagement by organising the *Visual Law workshop on the accessibility of privacy policies and consent forms and the respective requirements stemming from the Regulation 2016/679 on the protection of natural persons,* with regard to the processing of personal data[20] (further as the GDPR) and the Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (further as the e-Privacy Directive)[21] in M12. The Workshop gathered around 10 experts from academia, business, data protection authorities, consumer representatives (further as the speakers) as well as around 90 privacy experts and data subjects who had no

[18] PIMCity website: online access at pimcity-h2020.eu [accessed on 2020-11-11].
[19] Public Engagement in Responsible Research and Innovation. European Commission: online access at https://ec.europa.eu/programmes/horizon2020/node/766 [accessed on 2020-11-11].
[20] Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.
[21] 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) [2002] OJ L 201, 31.7.2002, p. 37–47.

prior knowledge of the topic (further as the participants). The speakers of the Workshop have discussed the EU legal requirements for privacy policies and consent forms and ways to improve the accessibility of these instruments. Besides, the Workshop has embodied an interactive session during which the participants provided extensive feedback on their experiences with these instruments, feedback on their accessibility as well as recommendations for improving it. PIMCity partners take into account the feedback received while designing the project solutions, drafting privacy policies and consent forms. Examples of the feedback are provided in the table below.

| **Feedback on accessibility of privacy policies and consent forms** | |
|---|---|
| **Questions** | **Answers** |
| *What makes privacy policy inaccessible for a user?* | 23% the document is too long<br>7% there is too much text<br>7% there are too many legal words<br>58% a bit of everything |
| *Do you think it is possible to make privacy policies more accessible (easy to be understand)?* | 86% yes<br>12% no<br>2% I do not know |
| *Rank the actions you consider necessary to improve privacy policies* | 1st shorten the text<br>2nd improve the format<br>3rd use less legal words<br>4th add icons |
| *Any additional suggestions on how to improve privacy policies and consent forms?* | "*Avoid verbose sentences keeping flat English messages when law-related topics are addressed ", "use more media and less words ", "give simple options – so that people can choose an understand ", "Less vague language ", "clearly describe the effect of the customer choices ", "svisualised cues, icons ", "design thinking* " |

KU Leuven – CiTiP also produced a White Paper in which key legal issues from both the data subject's and the data buyer's side is highlighted and commented on. The purpose of the White Paper is to provide an overview of voids in the legislation and to assess the evolution of case law, as well as potential solutions, if any.

Albeit less directly, KU Leuven – CiTiP also contributed to public engagement by disseminating its deliverables on the project's website. KU Leuven – CiTiP also contributed to public engagement by communicating about the project's activities through its social media channels (primarily Twitter) as well as through KU Leuven – CiTiP newsletter (internally). Individual researchers also contribute to public engagement by communicating

about the project's activities through their LinkedIn accounts. Particular details are provided in the deliverables of WP6.

**Telefónica Investigación y Desarrollo** (further as Telefónica) disseminates the generated scientific knowledge to the society via periodic events like the popular *Telefónica Innovation Day.*[22] During this event, which takes place every year, the Telefónica Group, S.A. shares with the public the company's strategy on how to bring disruptive solutions to the customers and how it plans to continue transforming the company. Additionally, Telefónica disseminates the research findings to the research community by giving talks and lectures, attending scientific panels and by organizing workshops, seminars, webinars and hackathons. The external communication channels of Telefónica include: (i) the corporate homepage; (ii) social networks: LinkedIn, Facebook, Twitter, Instagram, Flickr; (iii) video Channels: Youtube; (iv) press release deliveries; (v) RSS feed; (vi) Newsletter; (vii) Annual Reports; (viii) blog posts (Telefonica, BlogThinkBig); (ix) media. The team of Telefónica participating in PIMCity is committed to leverage such public engagement channels as above listed fully, to ensure the proper and wide dissemination of the research results and deliverables of PIMCity, both internally within the Telefónica Group and externally to the society.

**NEC Laboratories Europe** (further as NEC) ensures public engagement by developing different co-creation activities involving both citizens and civil organizations. To this end, the company has incorporated the design thinking [23] methodology that starts the Research and innovation with the direct involvement of the stakeholders. Moreover, the NEC X initiative (a startup incubator for the best research ideas) helps to bring the research ideas to a real product. This initiative has been designed to ensure citizens can enjoy the latest advances in the wide range of different technologies developed by the Research lab. For PIMCity, NEC disseminated the material generated through press releases and social networks via the researchers involved in the project.

**Politecnico di Torino** (further as POLITO) is an active member of national and international networks related to Public Engagement and RRI, such as APEnet (Atenei per il Public Engagement), the national working group on a gender-responsive budget of the Conference of Italian University Rectors and the CESAER task forces on Open Science and Human Resources. POLITO also has a specific office devoted to the press and public relation (Ufficio promozione e imagine – https://www.politocomunica.polito.it/); their tasks are focused in promoting dissemination activities, organizing events for the public and communicating the projects active on the territory, involving public and private entities as well as all citizens. POLITO has issued already two press releases: the first following the project kickoff meeting (https://www.pimcity-h2020.eu/event/kick-off-meeting-in-turin/) and the second following the release of the PDK (https://poliflash.polito.it/en/research_innovation/new_software_for_personal_data_platforms ).

In the light of PIMCity, POLITO collaborates in editing the PIMCity newsletter in coordination with other partners of the project. POLITO's teams are involved in circulating the news related to PIMCity and support the creation of the content published on social media (managed by the partner AUI). POLITO collaborates in organizing the workshops in order to involve academic and general public. In particular, during the demonstration activities foreseen for the future steps of the project, POLITO aims to involve students and researchers during the public events organized by POLITO (e.g., *Notte dei ricercatori, Biennale Tecnologia*).

[22] Telefónica Innovation Day: online access at https://innovationday.telefonica.com/ [accessed on 2020-11-11].

[23] Design thinking. Wikipedia: online access at https://en.wikipedia.org/wiki/Design_thinking https://ec.europa.eu/programmes/horizon2020/node/1031 [accessed on 2020-11-11].

The **University Carlos III of Madrid** (further as UC3M) defines its internal regulation regarding RRI in its *Code of Good Research and Transfer Practices of the Carlos III University of Madrid*,[24] approved in December 2017 (in Spanish).

The research groups at UC3M count on the support of the UC3M institutional communication structure to help them to achieve their objectives of information transfer to the general public regarding their initiatives, processes and results. These activities are undertaken in the area of scientific information and knowledge, through the Office of Scientific Information (further as OIC in Spanish; Oficina de Información Científica) from the Vice-Chancellor's Office for Communication and Culture of the UC3M. This unit is linked to the *Autonomous Community of Madrid's and the Network of Scientific Culture and Innovation Units* (UCC+i) of the *Spanish Foundation for Science and Technology* (FECYT – *Fundación Española para la Ciencia y la Tecnología*). The OIC uses a variety of formats and journalistic genres to emit scientific information regarding research projects and their findings (in Open Access), along with other subjects of interest in the area of Research and innovation.

It likewise participates in outreach activities such as *Science Week, Fairs of Science or European Researchers' Night* to disseminate scientific advances to society at large. These science education activities seek public engagement through interactions with the population, promoting a 'science with and for society' philosophy. In addition, the OIC creates and emits piece of news, press releases, videos and other multimedia products weekly about scientific results. In each communication campaign, the OIC prepares a dossier to track the impact that these contents have in webs and communication media. Diffusion Communication Channels: (i) Notiweb. Madrid+d; (ii) SINC Agency; (iii) AlphaGalileo; (iv) EurekAlert! & EurekAlert! Chinese. UC3M's external communication channels: (i) news on the corporate homepage: www.uc3m.es; (ii) social networks (LinkedIn, Facebook, Twitter, Instagram, Weibo); (iii) video channels (Youtube, UC3M Media); (iv) press release deliveries; (v) electronic newsletters for specific audiences; (vi) UC3M Innovation and entrepreneurship newsletter UC3M; (vii) images. UC3M's internal communication channels: (i) UC3M newsletter; (ii) TV Digital Signage; (iii) Corporate Magazine UC3M. Also project deliverables, media coverage and public publications.

In the light of PIMCity, the UC3M's team leverages the UC3M's OIC to disseminate the results of PIMCity both publicly internally at UC3M and externally to the society. Moreover, the UC3M's team participating in PIMCity has an established record of appearances in mass media to disseminate their research results as well as share their opinions as international recognized experts in different topics such as privacy, big data and technology. Finally, the UC3M team also develops its public engagement activity through the participation and release of PIMCity deliverable and research papers.

**IMDEA Networks** is active in dissemination activities. During PIMCity and for each communication campaign the impact that the published contents have over web audiences and communication media is measured and then registered in a news clipping, likewise for PIMCity related publications in the website site/blog about workshops, events or any other sort of items we promote as part of the PIMCity project.

Additionally, IMDEA Networks has several generic Diffusion Communication Channels to reach the public and disseminate the results of our research with reference to our open access platform:

- Notiweb. Madrid+d

[24] The University Carlos III of Madrid: https://e-archivo.uc3m.es/handle/10016/26071 [accessed on 2020-11-11].

- AlphaGalileo (international/bilingual) (https://www.alphagalileo.org/en-gb/)
- DiCYT (international/bilingual) (http://www.dicyt.com/espana):
- EurekAlert! (international/bilingual) (https://www.eurekalert.org/):
- Globedia (national/Spanish) (http://es.globedia.com/)
- Sciencex.com (international/English) (https://sciencex.com/media/) (relays to phys.org or Tech Xplore: techxplore.com)
- Agencia SINC/FECYT (national/bilingual) (https://www.agenciasinc.es/)
- Total Telecom (international/bilingual) (https://www.totaltele.com/post-article/press-release)
- Blog "Sociedad de la Información" hosted by mi+d (national/Spanish)

IMDEA Networks' external communication channels are mainly: News on the corporate homepage: https://networks.imdea.org/; Social Networks (LinkedIn, Facebook, Twitter, Instagram); Video Channels (YouTube: IMDEA Networks Institute).

IMDEA Networks' internal communication channels are mainly: TV Digital Signage; Media Coverage; Public publications.

**IAB Spain** is committed to disseminating the generated scientific knowledge to society. This happens via periodic events like the Inspirational festival and the Congreso de Regulación where every year IAB Spain shares with the public, and especially with the online advertising ecosystem, new data and advertising strategies while fulfilling with all regulations, particularly data protection ones. Additionally, IAB Spain disseminates the research findings to the online advertising ecosystem by holding webinars, organising workshops and others. The external communication channels of IAB Spain include: (i)news on the corporate homepage; (ii)social networks (LinkedIn, Facebook, Twitter, Instagram); (iii) video channels (Youtube); (iv)IAB Spain Newsletter; studies, Reports and other documents.

IAB Spain is committed to leverage such public engagement channels as above listed fully, to ensure the proper and wide dissemination of the research results and deliverables of PIMCity, both internally within the members of IAB Spain and externally to the society. Also, part of the role of IAB Spain in PIMCity Project is getting feedback of the online advertising ecosystem and for that reason, organized two workshops involving these types of companies for disseminating results and for demonstration purposes.

**Association of Internet Users (AUI)** has been working since 1995 in collaboration with other agents to promote a good use of the Internet and for several years with special attention to safeguarding the rights of citizens in the digital environment and everything related to the right to privacy and intimacy.

AUI is an active member in the Internet Governance Forums (IGFSpain[25], EURODIG[26] and IGFglobal[27]) that organize annual meetings and where from AUI we have worked so that the PIMCity project and the European strategy on the use and exploitation of personal data would be present, achieving that in its annual editions had sessions and round tables dedicated to this topic in which different speakers of the PIMCity project as well as responsible of the European Institutions have participated.

We are currently working to create a Dynamic Coalition[28] within the IGFGlobal focused on the PIMS model, which places the user at the center of the personal data exploitation system and brings together all the parties interested in promoting this data exploitation model.

AUI actively participates in the activities organized by other agents whose focus is privacy or development around personal data such as MyData[29] or BDVA/DAIRO[30] where we have actively contributed to the organization of discussion seminars in which the PIMS model and the role of intermediaries has been at the center of the discussion.

AUI's commitment is also with the public administrations both at Spanish and European level to whom we send information about the project and to whom we invite them to the round tables and seminars we organize to help us and engage in the dissemination of the results. The Ministry of Economy and the Directorate General for Digitalization and Artificial Intelligence, the Belgian and Spanish Data Protection Agencies, the European Data Supervisor and DGCONNECT are some of the institutions that have collaborated in some of the seminars organized by AUI and related to PIMCity project.

AUI organizes activities related to the European Data Protection Day in Spain and Latin America on January 28th, and to the World Internet Day (May 17th) and in both, the right to Privacy and to the Protection of personal data is worked with different social groups (school children, elderly people, underprivileged groups, etc).

AUI is also committed to the active dissemination and diffusion of the results of the project and therefore all its communication channels – web and social networks – re recurrently echoing the different advances and stages that are being achieved in the project.

25 www.IGFSpain.org
2020_edition: www.jornadasigfspain.es/timetable/event/sesion-7-modelos-para-generar-beneficios-con-los-datos-en-la-ue-en-ingles/
2021_edition: https://jornadasigfspain.es/documentacion/Programa%20Jornadas%20IGF%20Spain%2016-17%20de%20noviembre.pdf

26 www.eurodig.org
https://eurodigwiki.org/wiki/Data_Sovereignty_and_Trusted_Online_Identity_%E2%80%93_COVID-19_Vaccination_Data_%E2%80%93_WS_03_2021

27 https://www.intgovforum.org/

28 https://www.intgovforum.org/en/content/dynamic-coalitions

29 online2020.mydata.org/programme/#event-27

30 Data Vaults - https://bit.ly/2XSpTFT Building the next generation persona data platforms - https://bit.ly/3eFYhe3 Brokerage and market platform - https://bit.ly/2U3sHiC

**Wibson** is committed to innovate, test, launch and validate new data products to be used for both citizens and companies. Even more, Wibson works to launch no-code solution to be used by SMEs and freelancers. Wibson participates constantly in events organized by MyData, abstartups in Brasil and also in its own workshops and conferences. Wibson shares anonymized statistics of the industry and also has a blog where they teach and educate people on best practices in the new age of data.

During 2021 Wibson launched 3 different solutions concerning data including a cookie consent manager tool, a DSR platform and a SaaS to automate Privacy Impact Assessments. Wibson always added a free plan to allow everyone to test, use and validate its innovative solutions.

Wibson communicated its findings, and educational information by Instagram, Linkedin, Twiiteer, Facebook, webinars, workshops, Youtube, newsletter and blog.

### 7.1.2. Open Access

PIMCity partners recognise that making research results more accessible *contributes to better and more efficient science, and to innovation in the public and private sectors*.[31] In relation to this, the default policy of the PIMCity project is to make its data publicly available through public copyright licenses (e.g. Creative Commons), archiving it on the publicly accessible project website and OPENAIRE compliant repositories. Most PIMCity outputs is available for auditing, re-use and verification following the Open Access principles. All publications are available from the project website (see https://www.pimcity-h2020.eu/dissemination/publications/) and on the Participant Portal continuous reporting tool offered by the EC. The complete list of publications can be found in D6.6 too.

*Input of different partners*

Multiple partners such as AUI, KU Leuven – CiTiP, NEC, POLITO and ERMES have already contributed to the Open Access by disseminating their deliverables on the project's website pimcity-h2020.eu.

**KU Leuven – CiTiP** underlines that KU Leuven strongly believes in Open Access as the way to increase access to knowledge. KU Leuven supports *Green Open Access* with the deposit obligation in the institutional repository Lirias. KU Leuven also supports Fair Open Access, a.k.a. non-profit Gold Open Access, with the KU Leuven Fund for Fair Open Access.[32] In the light of PIMCity, KU Leuven – CiTiP contributed to the Open Access by disseminating their deliverables on the project's website pimcity-h2020.eu.

**Telefónica Investigación y Desarrollo** is committed to Open Access of research results and other data, whenever possible and whenever it does not contradict its obligation and legitimate interests to protect the results. In particular, Telefónica is obliged to examine the possibility of protecting its results and must adequately protect them — for an appropriate period and with appropriate territorial coverage — if:

- (a) the results can reasonably be expected to be commercially or industrially exploited and

[31] Open Science (Open Access). European Commission: online access at https://ec.europa.eu/programmes/horizon2020/node/1031 [accessed on 2020-11-11].

[32] Open Access @ KU Leuven: online access at https://www.kuleuven.be/open-science/what-is-open-science/scholarly-publishing-and-open-access/open-access-kuleuven [accessed on 2020-11-23].

- (b) protecting them is possible, reasonable and justified (given the circumstances).

When deciding on protection, Telefónica must consider its own legitimate interests and the legitimate interests (especially commercial) of the other beneficiaries in the project. In relation to this, Telefónica is committed to following the FAIR Data management guidelines in H2020 EU projects, and whenever possible and applicable, use Open Access repositories for disseminating research results and academic articles, such as www.arxiv.org, www.researchgate.net, or generally open-access journals and conferences such as the ones under the Foundations of USENIX, ICLR, etc. The same principles apply for sharing data, by using Open Access dataset repositories such as www.openAIRE.eu, www.zenodo.org, or generally available online storage repositories, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible. Finally, Telefónica contributed to the Open Access of the project, by disseminating their deliverables on the project's website (i.e., www.pimcity-h2020.eu) and source code via project's account on Gitlab (i.e., https://gitlab.com/pimcity).

**NEC** fully endorses the Open Access policy marked by the European Commission for all the outputs generated during the course of publicly funded research. As such, NEC has published dozens of papers during the past years following either a Green or Gold Open Access standard. For PIMCity, NEC makes publicly available all the papers generated during the project. Typically, by publishing in conferences and journals that are open, and uploading it to personal websites when that is not possible. Moreover, NEC participated in the different workshops organized by the project.

**POLITO**. In 2018 the Rector of POLITO appointed for the first time a Rector's Advisor for Open Science and in December the Governing Bodies approved the *Politecnico di Torino Policy on Open Access to Scientific Publications*. In 2018 an inter-departmental working group was created to raise awareness about open science and Open Access; the working group involved staff of the *Library and Museum Department, of the Research Support Department and the Quality and Evaluation Division*. The Strategic Plan POLITO4IMPACT strongly encourages and promotes Open Access to research results in order to bring POLITO's policies closer to the level of international best practices. Moreover, several trainings and events are organized on a regular basis.

In the light of PIMCity, as part of Horizon 2020, POLITO ensures the free accessibility of every result. POLITO published the results of the Research in Open Access modality. As of November 2021, all POLITO's publications are *Green Open Access*, and the final versions of the publications are deposited into an open repository and are dependent on the funder or publisher. Furthermore, the research group commits to publish artefacts (data and source code) used during the development of the Research in open repositories, to allow reproducibility, and foster further studies. Additionally, as coordinator of PIMCity, POLITO ensures that all publications (articles, book chapters, thesis dissertations, papers, including but not limited) are published on the PIMCity website[33] and are freely available for the public, with each publication linked to the repository and with references to the authors.

At **UC3M**, the mandates and issues related with *Open Science,* are managed by the Library Service, which reports to the Vice-President for Strategy and Digital Education, who collaborates, for this purpose, with the Research Service. The Library offers advice about all aspects related to Open Access to UC3M researchers. It also undertakes monitoring of the mandates of the funding organisms through the deposited research results in the *e-Archive and e-scienceData institutional repositories*. The Library, within its search engine, integrates all the existing sources with Open Access content to be used by UC3M users. Similarly, the results are available at the UC3M-Research Portal. The legal framework that affects the *Open Science* policies at UC3M, are

[33] PIMCity: online access at www.PIMCity-h2020.eu/dissemination/publications/ [accessed on 2020-11-11].

conditioned by the fact that it receives public financing for Research from the regional government of Madrid, MINECO and the European Union.

In the light of PIMCity, the UC3M team commits to release publicly all data, software and code produced as part of PIMCity. The only exception to this general rule would be any data, software of code subject to be part of commercial activity of being patented. These mechanisms avoid the public release of the asset, especially in the case of patents.

In addition, the UC3M team commits to make publicly available all the scientific publications produced by its work in PIMCity following either the Gold or Green Open Access options.

**IMDEA Networks** has the support of the Communication & Operations department, also for the Data Transparency Group (DTG) in PIMCity, in order to help them achieve their objectives of results dissemination and publication to the general public regarding their initiatives, processes and results.
The Communication & Operations department uses a variety of professional formats and journalistic genres to emit scientific information regarding research projects and their findings (in open access), along with other subjects of interest in the area of research and innovation.
In addition, the Communication & Operations department creates and emits piece of news, press releases, videos and other multimedia products weekly about scientific results.
In this respect, our project outputs use widely accepted self-archiving repositories (e.g. IMDEA, recently upgraded to **https://dspace.networks.imdea.org/**) or trusted repository services for the research community such as arXiv, ResearchGate, Academia or Zenodo. IMDEA also has experience with releasing open-source software in the gitlab repositories indicated here for the project, deliverables of the project at the project site, and research publications at recent international conference venues in our public team webiste and the dspace repository as for a recent IMC paper. In case of gold Open Access, the article is immediately published in Open Access mode. Nevertheless, the peer-reviewed scientific publication (either the published version, or the final peer-reviewed manuscript accepted for publication) is still deposited in a trusted and public repository.

**ERMES** is compliant with Open Access policy marked by the European Commission. This holds for all results generated by publicly funded research initiatives. As such, since its establishment all scientific paper produced by ERMES have been published with either a Green or Gold Open Access standard. Also for PIMCity, ERMES makes publicly available all the documents and results generated during the project using public platform for sharing of code, data, reports and papers.

**IAB Spain** website (+20000 users' / month) has a restricted area, only for members, but also has an open access area. For disseminating research results and outputs of PIMCity Project IAB Spain uses the open-access part of its website so that any internet user has access to the project's outputs. IAB Spain is committed to open access to research results and other data and to follow the FAIR Data management guidelines in H2020 EU projects.

**Internet Users Association (AUI)** supports and is committed to the open access policy set by the European Commission for all results generated within publicly funded research projects.

Our Internal strategy sets out a vision for encouraging and leveraging the transformative, innovative and collaborative power of open source, its principles and development practices. It promotes the sharing and reuse of software solutions, knowledge and expertise, to deliver better services that benefit society and lower costs to that society. AUI commits to increasing its use of open source not only for European projects, but also in all areas where we need to incorporate new software.

**Wibson** believes that innovation nowadays is done and accomplished working in a collaborative way, that's why they support and encourage open access policy set by the European Commission for all results generated within publicly funded research projects.

Wibson communicates about their research in media, social media and during various webinars and workshops. For example, this year Wibson shown the impact of GDPR and CCPA to European and American companies when answering to a DSR (Data Subject Request). This analysis was done inhouse and using Wibson's technology. However, Wibson shares it to contribute to development of other companies and to contribute to societal awareness.

### 7.1.3. GENDER BALANCE

PIMCity partners recognise the importance of fostering gender balance in research teams and decision-making as well as of integrating the gender dimension in research and innovation content. PIMCity partners assume that ensuring gender balance shall help to improve the scientific quality and societal relevance of the PIMCity solutions.[34]

*Gender balance in research teams and decision-making*

PIMCity partners undertake various equality measures in order to ensure equal opportunities in relation to HR processes (e.g. recruitment, training, promotion, work-life balance) and to address gender-related biases and under-representation. Although there are more male researchers in the PIMCity working team,[35] this underrepresentation reflects the well-documented gender gap in computing occupations and STEM careers, inventions and scientific publishing.

*Integrating the gender dimension in research and innovation content*

PIMCity partners undertake proactive measures to ensure all research and innovation activities are carried out in a gender-sensitive way. PIMCity partners acknowledge gender in every aspect of the research formulation, methodology, and outcome planning, including:

- gender-based differences related to the perception of vulnerability in privacy threats: studies suggest that there are differences related to how men and women react to privacy threats. PIMCity partners shall make sure it explores, addresses and reports any possible gender and sex-based nuances in both the design and development phases;
- gender-based differences related to self-disclosure in the social web: women seem to hold more concerns about privacy and are less willing to share personal information in social web than men, particularly when it comes to revealing sensitive information publicly. PIMCity partners shall explore if similar gender-based differences apply when sharing information to third parties via trackers, for instance;
- gender-balanced composition of engaged groups (volunteers, interviewees, etc.): in planning qualitative Research that may include interviews, focus groups and discussions with experts, as well as building communities of early adopters and beta testers, PIMCity partners shall ensure gender-balanced compositions.

[34] Promoting Gender Equality in Research and Innovation. European Commission: online access at https://ec.europa.eu/programmes/horizon2020/node/797 [accessed on 2020-11-11].
[35] Particular details are provided below in this section.

*Input of different partners*

**KU Leuven – CiTiP** aims to ensure gender-balanced research teams by ensuring an equal opportunity policy and by actively recruiting female researchers, sharing the tasks and leadership roles equally, to the extent possible. In the light of PIMCity project, it is notable that 50% of the most involved KU Leuven – CiTiP researchers were female, and 50% were male throughout the entire duration of the project. Besides, a female professor, Prof. Peggy Valcke, supervises the project.

Generally, the Gender policy at KU Leuven is shaped by the Integrated Gender Equality Plan 202136. This plan provides a detailed overview of the gender policy and gender indicators at the university.

**Telefónica Investigación y Desarrollo** is part of the Telefónica group: one of the top 25 of the most diverse and inclusive companies in the world, according to the index published by Refinitiv D&I 2020, which recognizes the top 100 companies that stand out most for incorporating both aspects into their work culture. Through its established Global Diversity Council and its Diversity and Inclusion Policy, Telefónica Group seeks to guarantee equal opportunities for all employees regardless of their gender, origin, age, sexual orientation and identity, abilities or any other personal characteristic. Under the umbrella of this company strategy, tools have been implemented to favour gender equality in selection and promotion processes; training programs have been created for managers on unconscious biases and diversity has been included as a quantifiable and measurable metric for variable annual remuneration, among many other actions.

**NEC** ensures fairness among genders in the hiring process, fostering the gender balance in research teams. All of the research activities developed by the lab consider the possible gender issues and integrate the gender dimension in research and innovation content. Finally, NEC yearly organizes the Girls day. The Girls day is an activity to attract young females to consider an education in the MINT professions (MINT = Maths, Informatics, Nature Science, Technology). Those professions are –in Germany- typical male professions. Currently, the work area is introduced to a group of schoolgirls and sample success stories of women in this profession presented. For PIMCity, NEC is not taking any action different from those taken in all the projects. The NEC's team in PIMCity includes a female researcher. However, her presence is not to ensure the gender balance, but taking into account that the project topics are in line with her interests.

**POLITO**. In 2007, POLITO approved its first Institutional Gender Equality Plan. With the new Strategic Plan 2018-2024 POLITO4IMPACT, POLITO moved one step further and set the following objectives regarding gender balance of students and staff: to increase the number of female students (up to 35%) and to support the career development of female researchers (+ 50%). In 2018 by sharing and implementing the principles of the European Charter for Researchers, a new governance structure of Equality@POLITO has been created and new initiatives to monitor and govern equal opportunities for the various aspects of diversity, starting from gender diversity, are ready to be launched. At last, up to November 2021, POLITO adopted its gender budget that is the basis on which to build an ambitious Gender Equality Plan.

[36] For details see KU Leuven Integrated Gender Equality Plan, online access at https://www.kuleuven.be/diversiteit/diversity/gender-equality-plan/index.html [accessed on 2022-04-11].

In the light of PIMCity, POLITO fully commits to involve a gender-balanced personnel, following the general guidelines. The following table provides the number of people involved into PIMCity.[37]

| Personnel | **Male** | **Female** | Total |
|---|---|---|---|
| Administration | 1 | 3 | 4 |
| Researchers | 4 | 1 | 5 |
| Students | 7 | 2 | 9 |
| Total | 12 | 6 | 18 |
| % | **66%** | **33%** | 100% |

POLITO aims to increase the number of female students involved, which is now below the 35% (target at the global level): the total number of students involved in Pimcity project so far is 9 (7 males and 2 female), working specifically on research and development activities.

POLITO is planning to involve female students, especially at the later stages of the project, particularly during the demonstration activities and recruitment actions, by adopting the general guidelines of POLITO.

**UC3M**. Currently, in the UC3M there is a Vice-rectorate for Students, Social Responsibility and Equality, and an Equality Unit (EU), responsible for developing, implementing, monitoring and evaluating the Equality Plans of the University. These plans are part of the Strategic Plan of UC3M and establish the objectives in terms of promoting equal treatment and opportunities, as well as the strategies or measures to achieve them. On 30th October 2017, the Governing Board approved the II Gender Equality Plan at the UC3M (further The Plan), which main aspects are summarized as follows. The Plan consists of 52 measures applicable to the entire university community. These measures are divided into 4 Axes (E) of intervention with a common theme: E1) Sensitize, communicate and train on equality issues; E2) Access, promotion, career and working conditions; E3) Conciliation and co-responsibility; E4) Teaching and Research. Each measure has a specific body responsible and related 'best indicators' to achieve.

Measures within Research (within the fourth axis: E4):

Within the fourth axis of intervention, Teaching and Research, 8 specific measures have been established to promote and raise awareness about gender equality in Research:

*E4.1. Encourage the inclusion of the gender perspective in national and European research projects.*

*E4.2. Call for grants for the organization of congresses and workshops on gender issues.*

[37] Data of November 2020.

*E4.3. Promote gender research through the Pilar Azcárate research awards.*

*E4.4. Analyse the evolution of women's participation in Research and disseminate the results.*

*E4.5. Increase the visibility of the Research carried out by women at the UC3M.*

*E4.6. Promote the participation of UC3M researchers in Science Week.*

*E4.7. Promote agreements with public and private institutions for Research and the transfer of research results with a gender perspective.*

*E4.8. Promote the inclusion of gender equality aspects as objectives of the research activity.*

In the light of PIMCity, UC3M underlines that while the presence of women in engineering departments and project is significantly small, the UC3M team participating in PIMCity includes female members. The technical work conducted by UC3M team in WP3 aims at retrieving the value of audiences (users' profile). As part of this work UC3M considers gender as one of the parameters to be considered what may help to identify potential biases in the value of audiences.

**IMDEA Networks Institute** aims to increase the proportion of women and therefore qualified female applicants are explicitly encouraged to apply. Until a balanced ratio of men and women has been achieved at the institute, preference was given to women if applicants have similar qualifications.

IMDEA Networks Institute actively promotes diversity and equal opportunities.

Applicants are not to be discriminated against in personnel selection procedures on the grounds of gender, ethnicity, religion or ideology, age, sexual orientation (anti-discrimination). People with disabilities who have the relevant qualifications are expressly invited to apply.

During WP3, likewise UC3M, we considered the value of gender data as something important but also a possibly sensitive feature for data valuation. Besides, if time allows, the PIMCity consortium is already proposing anonymization techniques that ensure published data respects strict privacy regulation (GDPR), regarding data subjects and Personally Identifiable Information (PII).

**IAB Spain** has a very marked gender equality policy. Approximately 75% of the employees are female, including all the top positions of the company, including its general manager, its marketing director, and its legal department director. IAB Spain has an internal complaints system created explicitly for abuse, discrimination and related issues. The above mentioned implies equal opportunities for all employees regardless of their gender, origin, sexual orientation. The IAB Spain team that participates in PIMCity is formed by exactly 50% of each gender.

**Internet Users Association (AUI)**, through its Diversity and Inclusion Policy, seeks to guarantee equal opportunities for all its employees regardless of their gender, origin, age, sexual orientation and identity, abilities or any other personal characteristic.

**Wibson** empowers diversity and equal opportunities. Wibson encourages having a diverse office where people can work and express theirselves freely. Wibson makes workshops about gender equality and introduces their policy to every new employee.

### 7.1.4. ETHICS

PIMCity partners recognise the importance of ethics as *an integral part of Research from the beginning to the end*.[38] PIMCity partners are aware that ethical research conduct *implies the application of fundamental ethical principles and legislation to scientific Research in all possible domains of Research,* including the domains researched in the light of PIMCity project. Taking into account that one of the most common ethical issues include privacy and data protection issues and the particular activities within the project, PIMCity partners pay particular attention to the compliance with all of the relevant national, European, EU and international privacy and data protection requirements. Besides, PIMCity partners aim to ensure there is no breach of research integrity, i.e. no falsification, plagiarism or other research misconduct.

PIMCity partners pay particular attention to the *European Commission Guidelines on ethics and data protection*.[39] Besides, since PIMCity partners shall work on algorithms, the project pays particular attention to ethical requirements regarding the use of artificial intelligence. They take into account, among others, the *AI HLEG Ethics Guidelines for Trustworthy AI*[40] *and the Guidelines on Artificial Intelligence and Data Protection under the Council of Europe*.[41] Overall, the consortium works in line with these ethical values: (i) research should be designed, reviewed and undertaken to ensure integrity and quality; (ii) research staff and subjects must be informed fully about the purpose, methods and intended possible uses of the Research, what their participation in the Research entails and what risks, if any, are involved; (iii) the confidentiality of the information supplied by research subjects and the anonymity of respondents must be respected; (iv) research participants must participate in a voluntary way, free from any coercion, or risk; (v) harm to research participants must be avoided; (vi) the independence of Research must be clear, and any conflicts of interest

[38] Ethics. European Commission: online access at https://ec.europa.eu/programmes/horizon2020/node/767 [accessed on 2020-11-11].

[39] Ethics and Data Protection. European Commission: online access at https://ec.europa.eu/info/sites/info/files/5. h2020 ethics and data protection _0.pdf [accessed on 2020-11-11].

[40] Ethics Guidelines for Trustworthy AI. European Commission: online access at https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai [accessed on 2020-11-11].

[41] Guidelines on Artificial Intelligence and Data Protection. Council of Europe: online access at https://www.coe.int/en/web/artificial-intelligence/-/new-guidelines-on-artificial-intelligence-and-data-protection [accessed on 2020-11-11].

or partiality must be explicit. The project partners revealed their compliance with the ethical requirements through their individual inputs in detail.

As a part of ethics, PIMCity partners also undertake to fulfil all of the legal requirements stemming from the relevant legal frameworks as defined in the deliverables of WP7, including the implementation of internal organizational and technical measures. Besides, taking into account that the PIMCity project is expected to collect and/or generate at least four broad categories of data, the project partners have developed the data management plan (D7.1) that elaborates on the management of these categories of data (further as the DMP). Overall, it reflects the consortium's comprehensive approach and joint efforts towards making research data findable, accessible, interoperable and re-usable (further as FAIR).[42] As part of making research data FAIR, the DMP provides the information on the handling of research data during and after the end of the project. The DMP indicates what data was collected, processed and/or generated, which methodology and standards were applied, whether data was shared/made Open Access and how data is curated & preserved (including after the end of the project).[43] The current version of the DMP is accessible on the project's website (pimcity-h2020.eu).

The PIMCity partners have also appointed the PIMCity data protection officer team to act as a single point of contact for data subjects wishing to exercise their rights, following the provisions of the GDPR and Article 29 Working Party Guidelines on Data Protection Officers.[44] The PIMCity data protection officer team developed a roadmap with actions to be taken if a data subject sends a request to give a response within the timeframes provided in the relevant legal frameworks.

More details on data management, including details on various technical measures are provided in D7.1, D7.6, as well as in D9.2 and D9.3 (confidential reports).

*Input of different partners*

**KU Leuven – CiTiP** played a role in ensuring ethics as an integral part of the Research. In particular, KU Leuven – CiTiP provided detailed guidance on the relevant privacy and data protection requirements both through its deliverables and through the consultations via calls and emails routinely. Among other things, KU Leuven – CiTiP advised the project partners on the data protection impact assessment, provided project partners with the detailed guidelines for consent management, privacy policies, data processing and joint controllership agreements, advised on the appropriate implementation of data subject rights, drafted templates of informed consent forms and privacy policies that were adjusted by the partners on a case by case basis taking into account particular details. Albeit less directly, KU Leuven – CiTiP also contributed to ethical Research through its workshop as revealed in part 1 of this document (public engagement).

**Telefónica Investigación y Desarrollo** has published Business Principles where the presented ethical code helps every part of the company to act with integrity, commitment and transparency. In this way, Telefónica's

[42] The DMP is drafted in accordance with the Guidelines on FAIR Data Management in Horizon 2020 as of 26th July, 2016, as issued by the European Commission Directorate-General for Research & Innovation and already referenced above.

[43] European Commission Directorate-General for Research & Innovation. Guidelines on FAIR <…>, p. 4.

[44] Article 29 Working Party Guidelines on Data Protection Officers, adopted on 13 December 2016, last revised and adopted on 5 April 2017 https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612048, and endorsed by the European Data Protection Board which replaced the Article 29 Working Party on 25 May 2018.

vision can be actualized and be sustainable over time based on trust and legitimacy. Its commitment to ethics and integrity is key for its transformation process. As per the established Responsible Business Principles, the company's ethical code arises from three basic values: *integrity, commitment and transparency*: values that are essential in promoting the relationship of trust the company wants to sustain with its stakeholders.
Security and respect for data privacy constitute the foundations of Telefónica's business and are of paramount concern when designing new services. Telefónica is committed to complying with all regulations in force across all markets in which it operates.
Telefónica follows the GDPR and has appointed a Data Protection Officer (DPO), who reports directly to the Board of Directors of Telefónica Group. The DPO oversees the Governance Model of Data Protection within the Group. The DPO leads the Global DPO Office, which performs the functions regulated in article 39 of the GDPR, focusing on (a) the design and coordination of the privacy compliance program at the global and corporate levels, in accordance with the corresponding risk analyses, and without prejudice to the specialties of each company and jurisdiction; and (b) the supervision of the implementation of said program.

**NEC** has a mission is to create value for all members of society. Looking to a brighter future, NEC believes that, with technology and co-creation, digital solutions can address society's needs. To this end, NEC has developed the *NEC Way.*[45] The NEC Way is a common set of values that form the basis for how the entire NEC Group conducts itself. Within the NEC Way, the *Purpose* and *Principles* represents why and how as a company NEC conducts business, whilst the *Code of Values* and *Code of Conduct* embodies the values and behaviors that all members of the NEC Group must demonstrate. While implementing the NEC Way NEC aims to create social value, revealed in detail below.

*Purpose: NEC creates the social values of safety, security, fairness and efficiency to promote a more sustainable world where everyone has the chance to reach their full potential.*

*Principles: The Founding Spirit of "Better Products, Better Services" Uncompromising Integrity and Respect for Human Rights Relentless Pursuit of Innovation.*

*Code of Values: Look Outward. See the Future. Think Simply. Display Clear Strategy. Be Passionate. Follow through to the End. Move Fast. Never Miss an Opportunity. Encourage Openness. Stimulate the Growth of All.*

*Code of Conduct: In order to realize a sustainable society, the NEC Group contributes to solving our customers' issues and advancing the resolution of social issues through the power of ICT, and thereby continues promoting our values of safety, security, fairness and efficiency. To provide our customers and society with excellent value, we need to have "Integrity", or in other words, high ethical standards and sincerity, as a foundation for all of our actions. This Code of Conduct provides detailed guidelines for this "Integrity". As a member of the NEC Group, we promise to comply in good faith with this Code of Conduct so as to be connected with people all over the world, beyond boundaries, and to enable the solving of social issues and sustainable economic growth for the future. It is based in: (i) Basic Position; (ii) Respect for Human Rights; (iii) Environmental Preservation; (iv) Business Activities with Integrity; (v) Management of the Company's Assets and Information, Consultation and Report on Doubts and Concerns about Compliance.*[46]

[45] NEC: online access at https://www.nec.com/en/global/about/the-nec-way.html [accessed on 2020-11-23].
[46] Code of conduct. NEC: online access at https://www.nec.com/en/global/about/pdf/necway/nec_code_of_conduct.pdf [accessed on 2020-11-23].

NEC guarantees an ethical approach in all the Research and business with a special focus on the respect for the Human Rights, the Environmental Preservation and the Business Activities with Integrity. During PIMCity, NEC researchers follow the code of conduct.

**POLITO**. In 2012 POLITO adopted the *Code of Ethical Conduct of the University Community* with the purpose of making all university members aware of the ethical principles which give rise to their rights and obligations as members of the university community. In July 2019, the Academic Senate approved the *Research integrity at Politecnico di Torino*, that is the programmatic manifesto of the internationally recognized fundamental principles that POLITO embraces in conducting its Research. In June 2020 also the Regulation of Research Integrity was approved to promotes the dissemination of the principles of integrity in Research and manage any violations. POLITO provides increasing support in the field of research ethics. In June 2020 the Regulation of the *Research Ethics Committee* was adopted; the Committee has been appointed in November 2020.

In the field of data protection and privacy, POLITO applies the data protection principles and faces issues concerning the processing of personal data as established by the GDPR. Since 2018 POLITO has a registered Data Protection Officer (DPO). The DPO ensures compliance of research objectives, processes and practices with the applicable EU and national law. The DPO acts as a data management consultant to the projects with particular regard to the lawfulness, fairness and transparency of the processing, to the use of data for legitimate purposes and connected to the institutional activities of the POLITO, as indicated in art. 2 of the Statute of the University, in a manner relevant to the treatment, respecting the principles of data minimization, accuracy, storage limitation, integrity and confidentiality, accountability.

In the light of PIMCity, POLITO embraces all previous activities and aspect reported above, actively working in close collaboration with the university's bodies. Specifically, for the data protection declaration for the privacy policy of the PIMCity website (www.PIMCity-h2020.eu/) POLITO ensures to not share collected data with any third parties. In case of consent for the use of cookies for statistical purposes, POLITO adopted a strategy to guarantee that the data collected through the use of these cookies is processed in POLITO's premises, using the Matomo Analytics that is installed in the servers hosted in POLITO's premises. POLITO takes appropriate administrative, technical and organizational measures against unauthorized or unlawful processing of any data or its accidental loss, destruction or damage, access, disclosure or use.

**UC3M**. The Research Ethics Committee of UC3M was established in September 2014 to assess and to monitor ethical issues of the University's research activities. It oversees all activities undertaken both, on the University's premises using its facilities, or on behalf of the University, by all University staff and students engaged in a research project. University policy for approval, structure and appeals procedures are laid out in the University's Research Ethics Regulation, adopted on 27th April, 2017, to ensure that the Research conforms with general ethical principles and standards. Researchers who intend to appeal the decision of any *Research Ethics Committee* at UC3M should follow such procedures and obtain approval before the Research begins. This Regulation applies across all subject disciplines and areas of study, regardless of the funding source, private or public, national or international calls, specific research aspects supervised by this Committee covers: (i) Human and animal research; (ii) Personal data protection; (iii) Protection of Fundamental Rights of People; (iv) Cooperation with developing countries; (v) Defence and security issues. If a project needs any approval or ethical screening, it was conducted by the UC3M Ethics Committee, regarding any of these research aspects.

For projects that collect and/or process Personal Data, it is required to follow all the procedures required according to institutional, local, national and international regulations. In particular, UC3M must comply with the GDPR and the Data Protection Act 2018 (DPA Regulation (EU) 2018/1725). This Act requires organisations that collect and use personal data, be transparent with individuals about how their data will be managed. The Act also imposes responsibilities and requirements on any organisation that handles personal data.

Moreover, under the GDPR, the University is required to appoint a Data Protection Officer (DPO) for their compliance. UC3M has appointed a DPO to advise the University on data protection law. UC3M DPO's primary roles are to ensure that any processing of any personal data of its staff, customers, providers or any other individuals (also referred to as data subjects), regardless the method used, is carried out in compliance with the data protection legislation and monitoring its performance against it, in cooperation with the data protection authority (for the EU institutions and bodies, this is the EDPS). The DPO also keeps an inventory and analysis of the processing activities and is involved in handling questions and complaints. The DPO also advises the organisation about the interpretation or application of the data protection rules. Everyone who processes personal data on behalf of the University must ensure that they comply with the University's Data Protection Policy: UC3M Data Protection Regulation.

In the light of PIMCity, the UC3M team does not plan to conduct any activity that requires the collection of personal data or implies any ethical concern. Therefore, it is not required to launch any procedure to obtain validation of its activity. Of course, if during the execution of the process, any new activity requires such validation, UC3M follows the established procedure described above. As part of the work in PIMCTY, the UC3M team has participated in the elaboration of D7.1 and D7.2.

**IMDEA Networks** ensures ethical research in PIMCity again through its Ethical Research Board, Guidelines and Recommendations available at IMDEA Networks that include principles in regards to respect for persons, beneficence, justice and respect of law. We inform to subjects involved in our research experiments for their consent by following our Ethical Research Board (ERB) as well in order to minimize risks, protect users' rights and their respective welfare. ERB is a committee formed of senior professors and faculty that decides on the suitability and measures the experiment requires (if any). Besides that, there is a Data Protection Officer (DPO) supporting IMDEA Networks data management operations.

In the PIMCity context IMDEA does not plan to collect user datasets that require the approval of our ethics board or DPO for use in PIMCity. In all cases, IMDEA does not employ human subjects in the context of PIMCity and in most cases it employs public datasets or data from partners in the PIMCity consortium. If in doubt, IMDEA holds direct internal contact with our ERB and DPO through our group leader at the DTG, in case we need to consult regarding ethical research regarding our contributions to PIMCity. This is in line with the vision of IMDEA Networks for Responsible Research and Innovation.

**IAB Spain**, defines its mission and in its ethics on the Code of Ethics and Conduct. This ethical code helps IAB Spain to act with integrity, transparency, impartiality, compliance and legality with data protection and environmental matters, among others. IAB Spain commitment to ethics and integrity is key for its daily operations within their activity and members. As per the established Code of Ethics and Conduct, the company's ethical code arises from five basic values: (i) respect for legality; (ii) transparency; (iii) impartiality; (iv) integrity; (v) respecting others. These values are essential in promoting the relationship of trust that IAB Spain wants to sustain with its members and with the online advertising ecosystem. Security and respect for data privacy are key in IAB Spain daily operations but also is committed to complying with all regulations. IAB Spain follows the GDPR, and data protection national laws and all of its employees have signed a data protection code of conduct which defines the risks when processing personal data and how they should be handled. Each company that processes personal data on behalf of IAB Spain has signed a data protection agreement.

In the light of PIMCity, aside from personal data (names and emails) collected during workshops, IAB Spain does not plan to conduct any activity that requires the collection of personal data or implies any ethical concern. Therefore, it is not required to launch any data protection risk assessment or data protection impact

assessment. If during the execution of the process, any new activity requires such assessment, IAB Spain performed it.

**Internet Users Association (AUI)**, the good and ethical use of technology is one of the objectives for which AUIwas created and founded in 1995. Since then, we have worked to develop frameworks and forums for debate on what is right and what is wrong on issues such as the right to accurate information, the right to freedom of expression, the right to be connected or the right to privacy and data privacy.

We are members of the Internet Rights and Principles Dynamic Coalition (https://internetrightsandprinciples.org/) an open network of individuals and organizations based at the United Nations Internet Governance Forum (IGF) and committed to making people's fundamental rights effective in the online world as well. Within this coalition we have worked on the elaboration of "10 Internet Rights&Principles"[47], a document translated into 27 different languages.

At a more local level we have collaborated from AUI as experts in the elaboration of the "Spanish charter of digital rights"[48] that was presented in July 2021 and adopted by the Spanish Government for the development of its digital transformation policies.

UI defines its ethics in the Internal Code of Conduct that commits all employees to the principles of transparency, impartiality, non-discrimination, respect for legality and sustainability in terms of data protection and the environment, among others.

AUI's commitment to data protection is also reflected in this Code of Conduct. AUI follows the GDPR, and national data protection laws.

**Wibson** ensures ethical organization by implementing its Code of Conduct through all workers and players in it. As developers of data privacy solutions, Wibson not only works with top lawyers to ensure GDPR, CCPA and LGPD implementation but also creates and implements their own privacy tools in the company to comply with data regulations. Since its creation, Wibson believed that the users are the owners of their data and hence since the beginning Wibson empowers them to be able to control it and manage their personal information.

Wibson shares MyData values to ensure a more transparent and inclusive data ecosystem and believes that the values they use to create our products such as transparency and control, should be also implemented through all the organization. Hence Wibson promotes full transparency and show their employees that Wibson represents those values and way to work.

47 https://drive.google.com/file/d/1MKByykdwe1Om1y_J6vXWVyIkZ6kSAII9/view?usp=sharing

48 https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf

### 7.1.5. SCIENCE EDUCATION

PIMCity partners recognise the importance of formal and informal science education in the society.[49] In relation to this, PIMCity partners contribute to it in a number of ways, e.g. by developing trainings and educational materials as revealed in detail below.

AUI coordinates the activities (WP6) that shall contribute to science education. Overall, activities coordinate by AUI shall contribute to maximising the opportunities of adoption, increasing public acceptability and building new awareness and educational opportunities around personal data platforms.

PIMCity training strategy addresses three different target groups – potential end-users and stakeholders in the digital and data-centric businesses, society at large and University students.

Firstly, PIMCity delivers online video tutorials providing guidance for the use of the tools developed for potential end-users and stakeholders in the digital and data-centric businesses, so that they would be able to educate themselves about the product use and configuration proactively. These video tutorials are freely accessible through major social platforms as well as in the project website, including but not limited.

Secondly, PIMCity prepares educational materials and sessions for engaging citizens into project topics (e.g. privacy, data ownership, etc.) in a broader scope than the innovation carried out in the project. These are imparted/translated into the local language of the audiences.

Thirdly, academic partners produce training materials for courses. For example, UC3M, delivers training material that will become part of UC3M Masters courses (e.g., Big Data, Cybersecurity) and degrees (e.g., in Data Science). The deliverables of WP6 that reflect the achievement of these goals can be found on the project's website pimcity-h2020.eu.

Besides, PIMCity partners recognise that effective dissemination and communication contribute to science education. In relation to this, PIMCity partners assume that their dissemination and communication activities contribute to achieving their science education goals as well. As briefly mentioned in the part 1 of this document (public engagement), the reports of dissemination and communication are provided as

[49] Science Education. European Commission: online access at https://ec.europa.eu/programmes/horizon2020/node/795 [accessed on 2020-11-11].

deliverables of WP6 on the project's website pimcity-h2020.eu (see, e.g. D6.3 as completed on 30th November 2020 and D6.4 completed on 30 November 2021).

*Input of different partners*

**KU Leuven – CiTiP** contributed to science education through its workshop as revealed in part 1 of this document (public engagement). Also, albeit less directly, it also contributes to science education by disseminating its deliverables and content of the workshops on the project's website pimcity-h2020.eu and communicating about the project's activities through its social media channels (primarily Twitter).

**Telefónica Investigación y Desarrollo** believes that education is the most powerful tool for reducing inequality and building the foundations for sustainable growth. Telefónica is committed to educational quality as a vehicle for social transformation. Telefónica, and specifically Fundación Telefónica contributes to science education via various ongoing projects:

- *EnlightED: a global conference that brings together prestigious international experts in education, technology and innovation to encourage a debate about Education in the digital era.*
- *ProFuturo: an initiative inspired by Pope Francisco that offers a comprehensive education solution to enable teachers to continue to make progress in their professional work and to manage their classrooms, whilst improving learning.*
- *Code.org: A joined project with Code.org aiming to boost learning in Computer Science among younger people. Code.org offers its courses to educators at schools and to parents at home and is also an excellent option for young people starting on their own. In addition, any person, anywhere in the world can sorganise a 'Code Time' event with one-hour tutorials in more than 45 languages.*
- *STEAM Challenge: A challenge aimed at young people between 14 and 25 years of age. It is time to put our skills in adaptation, collaboration, reflection and creation into practice.*

In addition, Telefónica Investigación y Desarrollo contributes indirectly to science education by disseminating its deliverables on the project's website pimcity-h2020.eu, and communicating about the project's research results via the company's large number of external communication channels (see Telefónica Investigación y Desarrollo Public Engagement section above), as well as invited and regular lectures at universities, summer schools and workshops, and other similar science-related education activities.

**NEC** contributes to the science education in two key ways: the execution of scientific projects leading to academic publications and the interaction with dozens of universities around Europe. Moreover, NEC participates in several activities to promote science in Europe. In particular, it organizes a reception for students during the prestigious Heidelberg Laureate Forum.[50] In PIMCity, NEC collaborates with the different events organized by the universities in the consortium.

**POLITO** promotes science education through many channels. PIMCity project is promoted through all these initiatives, events and association, with the perspective to involve students and general public:

- POLITO organizes public and free initiatives such as the *"Festival della Tecnologia" and "Biennale Tecnologia" and the "European Researcher's Night".*

[50] Online access at https://www.heidelberg-laureate-forum.org/ [accessed on 2020-11-23].

- POLITO has a specific structure, specifically devoted to the master's organization (https://didattica.polito.it/master/home/it/home) giving the opportunity to start an internship with Private companies.
- POLITO with Politecnico di Milano form part of ASP – Alta Scuola Politecnica (https://www.asp-poli.it/) dedicated to the best students, who are invited to follow specific multidisciplinary projects, in which they may practice the process of envisioning, framing, planning and implementing innovation.
- The initiative Giovani Talenti (https://didattica.POLITO.it/percorso_giovani_talenti/home) has been created to support the best bachelor students in developing their potentiality.
- Teams of students are supported by POLITO to be involved in projects and research initiatives (https://didattica.POLITO.it/pls/portal30/sviluppo.ateam.elenco2?p_lang=IT)
- Many hackathons have been organized and were organized on specific topics (list of past events is available here: https://www.POLITO.it/search/?lang=it&q=hackatlhon)
- POLITO is one of the 400 members, part of *Accademia delle scienze* (www.accademiadellescienze.it), association focused in spreading scientific knowledge through congresses, conventions, seminars, Workshop and other dissemination activities, with the aim to involve not only student but also general public and citizens.
- POLITO offers specific master degrees in ICT technologies (Data Science Engineering, Communication and Computer Networks Engineering, ICT for Smart Society, Computer Engineering) where specific course on privacy and data protection are offered.
- POLITO has decided to make substantial investments in order to finance PhD programs, which are deemed as strategic in the University agenda and education (http://dottorato.POLITO.it/it/home).
- POLITO has funded the SmartData@PoliTO center, which focuses on data science, big data and machine learning. The center organises specific dissemination events such as SmartTalks and SmartSeminars that aim at fostering the usage of fair data too.

POLITO commits to use all the above channels to disseminate the culture of fair data usage, personal data protection, data monetization, and related problems and solutions within PIMCity. Specifically, POLITO participated in public events, prepare classes and short seminars with different level of details, involve students and the general public, also in the light of the demonstration activities within PIMCity.

**UC3M,** as an academic institution, has a double commitment to Science Education. On the one hand, as a Higher Education Institution, it participates in formal education with different degrees, master and doctoral programs ranging across different social sciences and engineering disciplines. On the other hand, as a Public Institution, it is committed to disseminating the scientific knowledge to the society (See UC3M Public Engagement section above).

In the light of PIMCity and in the context of formal education, UC3M team participating in PIMCITY worked on transferring and integrating some of the results and findings obtained in the project to teaching material to be potentially integrated with relevant degrees and masters such as the Master in Cybersecurity, the Degree on Computer Science or the Degree in Telematic Engineering. In the context of science education UC3M Team, in addition to the actions described in Section of UC3M Public Engagement, also actively participated in different divulgation forums such as *Data Beers talks*, *Week of Science*, etc.

**IMDEA Networks** participates in outreach activities such as Science Week, Fairs of Science or European Researchers' Night to disseminate scientific advances to society at large. These science education activities seek public engagement through interactions with the population, promoting a 'science with and for society' philosophy. In addition, the institute organizes weekly seminars alternately invited talks with presentations by internal researchers.

In the context of formal education IMDEA Networks DTG group participating in PIMCity works on transferring and integrating some of the results and findings obtained in the project to teaching material to be potentially integrated in relevant UC3M degrees and masters such as the Master in Cybersecurity, the Degree on Computer Science or the Degree in Telematic Engineering.

“PIMCity Context” has planned to actively participate in different divulgation forums such as “Big Data Things”, “Internet Governance Forum Spain 2020” and others, “Data Beers Madrid”, etc. For instance, IMDEA

has already spoken at the following even highlighted by the Communications & Operations department at the IMDEA Networks website: https://networks.imdea.org/how-to-value-effectively-the-data/

**ERMES** deeply believes in science education. Indeed, thanks to its tight connections with universities, ERMES has the possibility to attract students aiming at concluding their studies with thesis in cyber security matters. Second, ERMES participates in scientific projects leading to academic publications and sharing of code, data and methodologies for inspiring new research initiatives. Moreover, ERMES has developed an internal upskill program which leverages popular web-based education platforms to allow its employees to improve their background in multiple disciplines.

**IAB Spain** provides education and training to its members and to the public in general. Those trainings are focused on digital matters, especially online advertising but also on data protection and other regulations. The education and trainings are offered in both online and traditional formats, also IAB Spain offer education and trainings to companies for covering their ad-hoc needs. IAB Spain host courses that have ethics sessions included by default, among others, on: (i) digital marketing and digital creativity; (ii) data analysis, regulation and data protection, GDPR; (iii) RTB; (iv) addressable TV and online video; (v) mobile ecosystem; (vi) marketing of influencers.

In the light of PIMCity IAB Spain works on integrating some of the results and findings obtained in the PIMCity project to teaching material to be potentially integrated in relevant courses such as the Advanced Course of RTB & Data and the Advanced Course of Legal and business aspects of digital advertising.

**Internet Users Association (AUI),** At AUI we believe that knowledge and information are very powerful tools to empower users of new technologies and we are aware that often ordinary people use technology without being aware of the risks associated with their use. We do not have the structure to provide training but we do develop actions oriented to sensitize organizations that work with citizens, educational centers (primary, secondary, university and adult centers) and social organizations so that they commit to include training and information courses on a good use and safe use of these technologies.

To this end, AUI works at the local and Latin American level with an Internet Promotion Committee[51] comprising 61 non-profit organizations, which every year develops citizen participation initiatives to empower users and raise awareness among the different social agents of the importance of training citizens.

In this dissemination and information strategy, AUI takes advantage of events such as the International Data Protection or Privacy Day (January 28), Internet Security Day (February 8, https://www.saferinternetday.org/) or World Internet Day (May 17, www.diadeinternet.org).

In relation to the PIMCity project, our commitment is twofold, since we are responsible for preparing the communication and training materials and also for carrying out the project's dissemination strategy.

Therefore, AUI continues with the strategy of participating in the events on data, privacy or PIMS organized by the different sectorial agents (MyData, CPDP, BDVA/DAIRO, IAPP) in the governance forums (IGFs, EURODIG, IGFglobal), ICANN) proposing sessions, seminars and providing the materials and tools developed in the project.

51 https://diadeinternet.org/comitedeimpulso

AUI also prepared a set of training and dissemination materials so that educational centers can take advantage of them and address the issue of personal data at different levels of depth and with special attention to audio visual media to facilitate the active participation of citizens.

**Wibson** contributes to social education through all their social media and channels. They believe that today as there is a lack of financial education, there is also a lack of privacy education, especially in Latam. With Wibson they generate workshops, webinars, they communicate best practices in their social media such as Linkedin, Youtube, Twitter, Instagram, Facebook, newsletter and blog. They also participate in events and are very active in the media to reach as more people as possible to empower them to own their data.

## 7.2. SOCIETAL IMPACT

PIMCity partners underline that the Research within PIMCity project addresses documented societal needs. The **goal** of PIMCity is to advance solutions for **transparency and to empower users to make informed decisions when sharing personal data**, i.e., to provide users with the information and tools to take **control over personal data**. The rationale behind this goal is consistent with documented societal needs with regard to fundamental rights, in particular privacy and data protection. [52] Overall, PIMCity contributes to implementation of privacy-enhancing technologies and privacy-by-design principles. PIMCity contributes to societal values related to accountability and transparency of (data) markets, as well as to the social and territorial cohesion efforts conveyed in the Digital Single Market strategy.[53] Besides, through dissemination and communication activities, PIMCity contributes to the visibility of research funding and its importance to addressing societal challenges such as privacy and data protection affecting citizens in their daily lives.

PIMCity partners expect that their **research benefits**

(i) **end-users** such as individuals, since PIMCity provides tools for easily understanding the monetary value of personal data, and taking informed actions that determine privacy settings;

(ii) **digital and data-centric businesses and services**, since they may improve their products and services; and

(iii) **enforcing and monitoring organisations**, including **regulatory bodies**, since they may take advantage of PIMCity tools for improving their capacity for providing transparency and privacy auditing services for helping and/or enforcing compliance.

The Research does not have a negative impact on the rights and values of any subjects, does not affect disproportionately specific groups and does not unduly discriminate them.

[52] These needs have been extensively documented in consultations, independent reports, impact assessments, including but not limited, and they are at the core of recent and large policy and regulatory efforts that address data subject's right to privacy and transparency from different angles. Besides, PIMCity explores the specific needs, expectations and reluctances of users and stakeholders with respect to the project planned results, which include an important empowering component (particularly within WP6).

[53] European Commission: online access at https://ec.europa.eu/digital-single-market/en/shaping-digital-single-market [accessed on 2020-11-11].

Although the Research within PIMCity does not address threats to society directly, PIMCity solutions are designed in a way that their components contribute to user's privacy and data protection. Accordingly, this report contributes to **preventing privacy threats indirectly**. Within WP7, PIMCity monitors the RRI, including ethical challenges related to privacy and data protection, making sure that these elements are taken into account while developing the project's solutions. It shall contribute to societal trust in PIMCity's solutions and shall serve as key means to achieve societal impacts, in particular with regard preserving the rights to privacy and data protection, and inclusive, relevant and socially acceptable research.

At the end of the project PIMCity partners also **validated and demonstrated project outputs** in order to assess public acceptance and engagement, including attitudes between different socio-economic, demographic and gender groups. Building on these findings, along with relevant literature, PIMCity partners address barriers in public perception and acceptability within the dissemination and public communication activities reaching society at large. The methodology for user engagement and community building shall also contribute to the trustworthiness of the system.

## 7.3. CONCLUSION

The updated Report on RRI confirms that PIMCity partners committed to responsible research and innovation via particular elements of public engagement, Open Access, gender, ethics, science education. Taking into account the particular goals and research activities of the project, PIMCity partners paid particular attention to ethics as an integral part of research and, specifically, to the requirements of privacy and data protection. In the light of the actions undertaken by the consortium as a whole and by different partners individually, PIMCity partners expect the results of the project to be inclusive, socially acceptable and to provide durable societal impact.

**ANNEX A – DATA PROTECTION IMPACT ASSESSMENT**

**Updated version of the Data Protection Impact Assessment as originally included in D7.1.**

**Legal Basis**

According to Art. 35 of the Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data[54] (further as the GDPR), a data protection impact assessment (further as a DPIA) should be carried out when the processing activity of personal data is *'likely to result in a high risk for the rights and freedoms of natural persons'* (Art. 35(1) GDPR).[55] The methodology used to develop the DPIA is based not only the GDPR provisions, but also the guidelines provided by Article 29 Working Party (further as

[54] Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.

[55] Art.35(1) GDPR: *'Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks'*.

WP29).[56] To this end, WP29, substantiating GDPR provision lists 'nine criteria' that should be taken into account to establish whether a DPIA should be carried out.[57] The activities characterising the PIMCity project are likely to fall into the two listed criteria. In particular: *'Data processed on a large scale'*[58] and *'Matching or combining datasets'*.[59]

In line with the legal requirements provided in Art. 35 GDPR, and in compliance with the privacy and data protection accountability principle,[60] PIMCity partners have carried out the DPIA in the light of the PIMCity project.

Based on Art. 35(7) GDPR, a DPIA shall include:

- a description of the context and purposes of the processing of personal data (context);
- a justification of the necessity and proportionality of the processing operations in relation to the purposes (fundamental rights);
- an assessment of the risks to the rights and freedoms of data subjects that might be generated by the processing activities;
- an explanation of the measures envisaged to address the risks, including safeguards, security measures, and mechanisms to ensure the protection of personal data and to demonstrate compliance with the GDPR (mitigation measures).

**Methodology**

To achieve comparable results, PIMCity partners have opted for a common approach concerning tools and methodology for the execution of the DPIA.

First, KU Leuven – CiTiP has provided clarifications and guidelines on the DPIA, and how this should be developed. Consortium partners NEC Laboratories Europe, Politecnico di Torino, Fastweb, Asociación de

[56] WP29, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, WP248 rev.01, adopted on 4 April 2017 as last revised and adopted on 4 October 2017.

[57] Ibid, p.10

[58] Art. 35(3)(b) GDPR. Large scale of data processing is included among the examples that GDPR provides that require a DPIA. Nonetheless, the GDPR does not provide a definition of what constitutes large-scale, though recital 91 provides some guidance. Contrary, according to WP29 guidelines following factors should be considered to determine whether the processing is carried out on a large scale: *'a. the number of data subjects concerned, either as a specific number or as a proportion of the relevant population; b. the volume of data and/or the range of different data items being processed; c. the duration, or permanence, of the data processing activity; d. the geographical extent of the processing activity'*. WP29, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, WP248 rev.01, adopted on 4 April 2017 as last revised and adopted on 4 October 2017, p.10.

[59] Matching or combining criteria concerns for example processing activities originating from two or more data processing operations performed for different purposes and/or by different data controllers in a way that would exceed the reasonable expectations of the data subject. Such criteria is strictly related to the purpose limitation principle, therefore, specific attention should be paid in regard to the lawful basis for processing.

[60] Art. 5(2) GDPR: *'The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 ('accountability').'*

Usuarios de Internet, Universidad Carlos III de Madrid, Telefónica Investigación y Desarrollo, IMDEA Networks, (further as the PIMCity partners) have concluded that, given their roles and foreseen activities in the PIMCity project, they shall carry out a DPIA.

Second, partners whose activity fall into the GDPR scope of application, have carried out its respective DPIAs individually. Following the suggestion of KU Leuven – CiTiP, the partners have developed their DPIA through the software PIA, provided by the French Data Protection Authority 'CNIL' ('Commission Nationale de l'Informatique et des Libertés').[61] CNIL software has been chosen for two main reasons: it is developed by a national data protection authority, and include all requirements listed in Art. 35(7) GDPR.[62]

Third, the final results of the PIMCity partners individual DPIA have been shared within the consortium through the project repository made available in Microsoft 'Teams'.

Finally, the results provided by partners have been collected and summarised by KU Leuven – CiTiP and provided in the table below.

More details on various mitigation measures are provided in D9.2 (confidential), coordinated by Politecnico di Torino. More details on verification of the implementation are provided in D7.5.

[61] https://www.cnil.fr/en/privacy-impact-assessment-pia

[62] Erik Kamenjasevic, Elisabetta Biasin, Safecare - Deliverable 6.1 Legal and ethical inventory and in-depth analysis, December 2018

### 1. Initial PIMCity Data Protection Impact Assessment – Project Wide

| DPIA Sections | Task | Partners' Initial Assessment |
|---|---|---|
| **CONTEXT** | Overview of the processing activities and purposes under consideration | PIMCity partners have described the activities involving the processing of personal data and the purpose of such activities. They have also described the nature of the data processed.<br>Sample activities of PIMCity partners include data collection, data storage, data use, data combination, and erasure.<br><br>PIMCity partners have also described and how they intend to achieve the declared purpose (e.g. collects and processes data from websites using automatic web scrapers to generate privacy metrics). |
| **FUNDAMENTAL RIGHTS** | An evaluation of the necessity and proportionality of the processing operations in relation to the purposes | PIMCity partners, according to their role within the project have declared that the processing purpose of their activity is explicit and legitimate.<br><br>Concerning lawful basis for processing personal data, most of the PIMCity partners have affirmed that consent is going to be the legal ground used for their processing activities.<br><br>Besides, PIMCity partners have also provided information about their activities regarding data accuracy, data minimisation, and data storage. |
| **RISK ASSESSMENT** | Assessment of the risks to the rights and freedoms of data subjects resulting from the processing activities | PIMCity partners have assessed the risks to the rights and freedoms of data subjects link to their activities. According to PIMCity partners risks can be generated by:<br>(I) Major Event<br>(II) System bug<br>(III) Third-party unauthorised access<br>(IV) Human error<br>Most of the assessments show a negligible risk.<br>A minor part of these assessments find a limited risk. |
| **MITIGATION MEASURES** | The mitigation measures envisaged for addressing the risks[63] | PIMCity partners have already developed a list of security measures and protocols to address the highlighted risks, taking into account the state-of-the-art of such security protocols.<br>In particular, PIMCity partners mitigation measures include:<br>(I) Physical access control,<br>(II) Logical access control,<br>(III) Hardware access control.<br>In particular, the security of personal data partners is ensured by encryption protocols and when necessary anonymisation functions. |

[63] Erik Kamenjasevic, Elisabetta Biasin, *Safecare - Deliverable 6.1 Legal and ethical inventory and in-depth analysis*, December 2018.

## 2. Polito

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | POLITO |
| CONTEXT | OVERVIEW | What is the processing under consideration? | Processing of web page content to generate privacy tags that summarize the usage of personal data of a website. This entails only the processing of web pages and no personal data |
| | | What are the responsibilities linked to the processing? *(data owner/data ubjects/controller/processor)* | POLITO will be controller for this activity, and the project. |
| | | Are there standards applicable to the processing? | There are no standards to apply, but general best practices are considered: data is maintained encrypted and transferred using secure channels. Data is accessible and processed by authorised personnel only, possibly using automatic procedures. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | We will collect web pages by automatically crawling websites. Web pages with content and automatic identification of third-party as tracker or advertisers |
| | | How does the life cycle of data and processes work? | By modelling the objects that are part of a webpage, and the servers distributing it as a graph, we will compute privacy tags using machine learning approaches. We will consider both supervised and unsupervised machine learning to automatically classify pages and websites according to the amount of personal data they collect. Data is collected and stored on a Hadoop cluster for processing with ML algorithms. |
| | | What are the data supporting assets? | Data will be stored on a Hadoop cluster hosted in POLITO premises and processed using big data platforms like Apache Spark running locally. |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | The goal of the data collection and processing is provide privacy tags. This is an explicit and legitimate goal in the context of the project execution. |
| | | What is the storage duration of the data? | We need to keep historical view of data to observe how the ecosystem changes over time. We plan to use long term storage to allow several implementation of the algorithms, validation of results, and eventually sharing of the data for reproducibility purposes. |
| | | What are the legal basis making the processing lawful? | We do not plan to collect and use any personal data. The collection of web pages will consider eventual restrictions imposed by the website, e.g., respecting the "robots.txt" constraints. |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | The data we collect is openly accessible on the web (i.e., web pages that can be accessed freely online by anyone).<br><br>While not all part of the webpage is necessary to implement the algorithms, the data collection will be minimized to the sole scope of the analysis. |
| | Fairness | Are the data accurate and kept up to date? | Yes |
| | | In case you rely on consent, is it really free? | No consent |
| | | How do you document the that people gave it? | N.a. |
| | | How can they revoke their consent? | N.a. |
| | | Could this generate chilling effects? | N.a. |
| | | Could this lead to discrimination? | N.a. |

| | | | |
|---|---|---|---|
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | N.a. |
| | Transparency | How will you tell people about your processing? | No personal data involved |
| | | How do you make sure the information reaches the persons affected? | As described in deliverable D7.5 |
| | | Is the information you provide complete and easy to understand? | Yes, as described in deliverable D7.5 |
| | | Is it (the provided information) targeted to the audience? | Yes, as described in deliverable D7.5 |
| | | In case you defer informing people, how do you justify this? | N/A |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes, as described in deliverable D7.5 |
| | | Are all purposes compatible with the initial purpose? | Yes, as described in deliverable D7.5 |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | No. |
| | | How can you ensure that data are only used for their defined purposes? | No other purpose possible. |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes |
| | | Do the data you collect measure what you intend to measure? | Yes, as described in deliverable D7.5 |
| | | Are there data items you could remove without compromising the purpose of the process? | No, as described in deliverable D7.5 |
| | | Do you clearly distinguish between mandatory and optional items in forms? | Yes, as described in deliverable D7.5 |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | By applying state of the art privacy preserving analytics. |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | No personal data collected |
| | | How do you ensure that the data you collect yourself are accurate? | Provided by the user itself. |
| | | How do you ensure that data you obtain from third parties are accurate? | N/A. |

| | | | |
|---|---|---|---|
| | | Do your tools allow updating / correcting data where necessary? | Yes |
| | | Do your tools allow consistency checks? | No |
| | Storage Limitation | Does EU legislation define storage periods for your process? | no |
| | | How long do you need to keep which data? For which purpose(s)? | Until 3 years after the end of the project (i.e. August 2025) for legal reasons. |
| | | Can you distinguish storage periods for different parts of the data? | No. |
| | | If you cannot delete the data just yet, can you restrict access to it? | Yes. |
| | | Will your tools allow automated erasure at the end of the storage period? | Yes. |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | Yes, the data will be stored in servers which are protected with start-of-the-art security techniques implemented by the POLITO IT & Security department including firewall, traffic monitoring, server access control through users and passwords, etc. |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | no |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | Yes |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | Yes |
| RISKS | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | No risks since we are not dealing with data subjects. |
| | | What are the main threats that could lead to the risk? | Some unauthorized person can access the data on the cluster storing the web crawling data. |
| | | What are the risk sources? | Possibly malicious attackers can access the datacentre and the data. Eventual some other internal people can accidentally access/delete the data. |
| | | Which of the identified planned controls contribute to addressing the risk? | Data is stored on the Big Data cluster - which is installed in the POLITO datacentre with strict access control. Access to the cluster is granted only to authorized users. Only people working on the project belong to the group of people authorised to access the data. The cluster hosting the data is hosted in the Polito datacentre, with air-condition, fire control, restricted access |

## 3. AUI

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | AUI |
| CONTEXT | OVERVIEW | What is the processing under consideration? | The participation of voluntary end-users to fill in surveys and to test PIMcity applications as beta testers through the web www.pimcity.eu |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | AUI is configured as data owner and processor. AUI' internal contact is Miguel Pérez Subias, Presidente and legal representative of AUI |
| | | Are there standards applicable to the processing? | There are no standards to apply, but general best practices are considered: data is maintained encrypted and transferred using secure channels. Data is accessible and processed by authorised personnel only, possibly using automatic procedures. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | The data is collected through forms on the website. Once the data arrives at AUI´s web server, the identification data is stored in encrypted form in a table and the rest of the data is disaggregated and stored in order to be able to work with it for the purpose of making calculations and statistics. |
| | | How does the life cycle of data and processes work? | The data remains on the AUI server until one years after the end of the project when all the data collected is destroyed. It is also possible to destroy the data of a particular user if he requests it through the channels enabled for this purpose. |
| | | What are the data supporting assets? | The data will be stored in a database locally, processed and available only for contact during the project with the participants in the different activities and are not given to third parties.<br>All servers for data collection are based on operating system Linux and configured to automatically install the latest security patches. Data are transferred using secure HTTPS channels and stored encrypted either on the local disk. |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | Yes. The aim of collecting this data is to be able to contact users who want to contribute to carrying out surveys or testing the results aimed at users. This is an explicit and legitimate goal in the context of the project execution. |
| | | What is the storage duration of the data? | One year after the end of the project |
| | | What are the legal basis making the processing lawful? | Those established by the GDPR. Each form informs you of the purpose for which your data is collected, how to exercise your rights and finally asks for the **consent** of each user. |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | Yes, we collect only the data necessary for the goal of the project. |
| | Fairness | Are the data accurate and kept up to date? | Yes, the registered user can at any time review and update their data. |
| | | In case you rely on consent, is it really free? | Yes, the user has all the information and can delete their data and revoke consent at any time. |
| | | How do you document the that people gave it? | Through a dated record in the server database. |
| | | How can they revoke their consent? | By email, postal or web form |
| | | Could this generate chilling effects? | No |

| | | | |
|---|---|---|---|
| | | Could this lead to discrimination? | No |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Yes, through the website (private area, contact form), by email and by post. |
| | Transparency | How will you tell people about your processing? | When you fill in the form, you are informed in a clear, simple and explicit way about the purpose, the data stored and how to exercise your rights. |
| | | How do you make sure the information reaches the persons affected? | A confirmation is requested via email where you are again informed of the data that is collected, purpose and how to exercise your rights. |
| | | Is the information you provide complete and easy to understand? | Yes |
| | | Is it (the provided information) targeted to the audience? | Yes |
| | | In case you defer informing people, how do you justify this? | Only those users who confirm their registration will be registered, otherwise the data will be deleted. |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes |
| | | Are all purposes compatible with the initial purpose? | Yes |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | No |
| | | How can you ensure that data are only used for their defined purposes? | The databases and files where this data is stored are specific to the PIMCity project and are not shared with other departments, projects or staff of the Association. |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | Not applicable since the data are only for this project. |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes |
| | | Do the data you collect measure what you intend to measure? | Yes |
| | | Are there data items you could remove without compromising the purpose of the process? | No |
| | | Do you clearly distinguish between mandatory and optional items in forms? | All information is mandatory since we only ask for the data we need for the project to work. |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | This use is not intended |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | None |
| | | How do you ensure that the data you collect yourself are accurate? | You check that the email provided by the user is functional, this is the only check you make. |
| | | How do you ensure that data you obtain from third parties are accurate? | It is not checked |
| | | Do your tools allow updating / correcting data where necessary? | Yes, via the web, by email and by post |

| | | | |
|---|---|---|---|
| | | Do your tools allow consistency checks? | No |
| | Storage Limitation | Does EU legislation define storage periods for your process? | No |
| | | How long do you need to keep which data? For which purpose(s)? | Up to one year after the end of the project in case you need to do any tests with the btatesters after the end of the project. |
| | | Can you distinguish storage periods for different parts of the data? | Not all data is deleted at the same time, unless there is an individual request which is handled immediately. |
| | | If you cannot delete the data just yet, can you restrict access to it? | Yes |
| | | Will your tools allow automated erasure at the end of the storage period? | Yes |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | Yes |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | Yes |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | Yes |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | Yes |
| RISKS | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | Use of the collected data for another purpose. |
| | | What are the main threats that could lead to the risk? | Impersonation to target third parties by impersonating the user whose data has been collected. |
| | | What are the risk sources? | Access to the database of the servers where these data are stored. |
| | | Which of the identified planned controls contribute to addressing the risk? | Unique own accounts are registered to detect if someone is making use of the database. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Very unlikely as the data is stored encrypted in the database and access to the database does not allow access to the collected data in a readable form. |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Very low |
| | | Planned or Existing Mitigation measures to the foreseen risks? | Contact with those affected so that they are aware of the fact and can make changes to update their accounts. |
| | Unwanted modification of data | What could be the main impacts on the data subjects if the risk were to occur? | None |
| | | What are the main threats that could lead to the risk? | None |
| | | What are the risk sources? | Access to the database of the servers where these data are stored. |
| | | Which of the identified planned controls contribute to addressing the risk? | Log and control access to servers databases |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited. |
| | | Planned or Existing Mitigation measures to the foreseen risks? | Database backups |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | None |
| | | What are the main threats that could lead to the risk? | Limited |
| | | What are the risk sources? | Access to the database of the servers where these data are stored. |
| | | Which of the identified planned controls contribute to addressing the risk? | Log and control access to servers databases |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |
| | | Planned or Existing Mitigation measures to the foreseen risks? | Database backups |

## 4. Ermes

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | ECS |
| CONTEXT | OVERVIEW | What is the processing under consideration? | ECS collects and processes data from websites using automatic web scrapers to generate Privacy Metrics, i.e., sets of scores and metrics to allow users understand which personal data is collected and how from websites and third parties connected to them. |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | ECS is configured as data owner and processor. ECS' internal contacts are Stefano Traverso, Head of Research, and Hassan Metwalley, CEO and legal representative of ECS. |
| | | Are there standards applicable to the processing? | There are no standards to apply, but general best practices are considered: data is maintained encrypted and transferred using secure channels. Data is accessible and processed by authorized personnel only, possibly using automatic procedures. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | ECS' web scrapers download all code files needed to render the page (e.g., HTML and Javascript files) as well as logs to API calls executed by the browser when rendering the page. |
| | | How does the life cycle of data and processes work? | For each website visited by a web scraper, we download all files and logs (see description above) and store them in the cloud in encrypted containers (e.g., AWS S3 buckets) which are accessible to authorized personnel only. Once collected, data are processed using BigData- and ML-based automatic tools which seek for patterns typically used by web services to collect personal data. The results are then stored in another encrypted container, still accessible to authorized personnel only. |
| | | What are the data supporting assets? | All servers for data collection are based on operating system Linux CentOS 7 and configured to automatically install the latest security patches. Data are transferred using secure HTTPS channels and stored encrypted either on the local disk or in cloud containers. Servers used for data analysis are based on on-demand AMI Linux. Data are copied on such servers for the time needed by the analysis. Once the analysis is completed, results are transferred to cloud containers (Amazon Web Services S3) using HTTPS secure channels. |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | The purpose of processing is specific as exclusively aimed at generating Privacy Metrics, and consequently provide the users with tools to understand privacy risks connected to given web services. Data are not used for other purposes. The purpose is explicit as the processing is properly documented and supported by the project. The purpose is legitimate as the output of our analysis depicts in a transparent way how web services actually collect and use users' personal data. |
| | | What is the storage duration of the data? | ECS stores data only for the time needed to provide the services based on Privacy Metrics. |
| | | What are the legal basis making the processing lawful? | ECS collects and treat data which are publicly available, results are representative of how personal data is used in the public web, and this processing does not require ECS to process personal data. |
| | | Are the data collected adequate, relevant and limited to what is necessary | Yes, we collect only data which are adequate and relevant for the goal of the project, and only in the amount |

| | | | |
|---|---|---|---|
| | | in relation to the purposes for which they are processed ('data minimisation')? | necessary to conduct the processing and provide a reliable and accurate output |
| | Fairness | Are the data accurate and kept up to date? | ECS collects data periodically to provide up-to-date results, and guarantees data accuracy by periodically inspecting them using ad hoc procedures. |
| | | In case you rely on consent, is it really free? | Data processed by ECS for Privacy Metrics is either public or provided by the company itself. Consent is really free. |
| | | How do you document that people gave it? | The data buyer will be provided with an interface (ticking a box) to give consent to data processing. |
| | | How can they revoke their consent? | The company/data buyer providing data will be able to revoke consent at any time by accessing the data buyer web interface and removing her data. |
| | | Could this generate chilling effects? | No. |
| | | Could this lead to discrimination? | No. |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Yes, by leveraging the web interface, the data buyer will be provided the means to exercise his rights. |
| | Transparency | How will you tell people about your processing? | Data buyers will be given the information about the processing in the web interface used to provide the data. |
| | | How do you make sure the information reaches the persons affected? | The interface will include a ticking box to ensure the information has been read. |
| | | Is the information you provide complete and easy to understand? | Yes, the interface is undergoing proper tests to check presentation of information is quick and easy to understand. |
| | | Is it (the provided information) targeted to the audience? | Yes, only authorized roles corresponding to authenticated parties participating PIMCity will be able to access the information. |
| | | In case you defer informing people, how do you justify this? | ECS operates to always provide information. Deferring may happen for technical reasons which will be communicated to parties involved in the processing. |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes, ECS has clearly defined the purposes of the processing. |
| | | Are all purposes compatible with the initial purpose? | Yes, ECS has not identified other purposes different from the initial one. |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | Yes, but ECS has defined proper policies to avoid data can be reused for other purposes. |
| | | How can you ensure that data are only used for their defined purposes? | Only authorized roles corresponding to authenticated parties participating PIMCity can access the data, and information about accesses (time, id of the party and data resource are logged). |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | We do not plan to make data for these purposes at the moment. |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes, they are. |
| | | Do the data you collect measure what you intend to measure? | Yes, they do, and they are limited to that specific purpose. |
| | | Are there data items you could remove without compromising the purpose of the process? | No, all data items are needed. |

| | | | |
|---|---|---|---|
| | | Do you clearly distinguish between mandatory and optional items in forms? | Yes, ECS will define which data are mandatory to provide for the purpose in the web interface made available to data buyers. |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | ECS does not plan to keep information for statistical purposes. |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | Information obtained by Privacy Metrics might be unhelpful for the users of EasyPIMS. |
| | | How do you ensure that the data you collect yourself are accurate? | It is in the interest of the data buyer to provide accurate data. |
| | | How do you ensure that data you obtain from third parties are accurate? | ECS does not collect personal data from third parties and does not plan to do that in the future. |
| | | Do your tools allow updating / correcting data where necessary? | Yes, the web interface will allow data buyers to correct and update information any time. |
| | | Do your tools allow consistency checks? | The specific nature of the data collected by ECS for Privacy Metrics (mail addresses of company's personnel) does not require to perform consistency checks. |
| | Storage Limitation | Does EU legislation define storage periods for your process? | ECS follows the general rule of storing the data for time needed to provide the service |
| | | How long do you need to keep which data? For which purpose(s)? | ECS stores data only for the time needed to provide the services based on Privacy Metrics. |
| | | Can you distinguish storage periods for different parts of the data? | No, the only personal data ECS might acquire are mail addresses. |
| | | If you cannot delete the data just yet, can you restrict access to it? | ECS can delete the data any time, but access to data is restricted by design. |
| | | Will your tools allow automated erasure at the end of the storage period? | No, ECS has not automated this part of the processing. |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | ECS will soon be certified with ISO 27001 standard for information security. Security risk assessment are conducted periodically as a certification requirement.<br><br>ECS stores data in servers and containers which are protected with state-of-the-art security techniques and technologies. These are:<br>• Cryptography: all data are stored and transferred encrypted.<br>• Access control: access to data is provided only to authorized personnel using RBAC policies and from authorized workstations using ad-hoc firewall policies. Physical access is provided to authorized personnel only.<br>• Traceability: access to data events are tracked. Access logs are stored for a maximum period of 1 month.<br>• Data partitioning: data are stored using spatial (e.g. per website) and temporal (e.g., per day) partitioning.<br>• Storage: data integrity is guaranteed by using cryptography techniques.<br>• Minimization of personal data: ECS does not collect personal data in the context of this project. |

| | | | |
|---|---|---|---|
| | | | • Website security: ECS follows ANSSI guidelines to guarantee security of its websites.<br>• Backup: ECS stores data to process in the cloud, where the cloud provider offer transparent backup policies. Data stored on physical servers are duplicated periodically with full and incremental backups and these are copied to the cloud.<br>• Network security: ECS uses network-level segmentation and filtering techniques to guarantee access from given workstations or IP addresses and isolate from the Internet both physical and virtual servers. No server in both cases are publicly accessible.<br>• Physical access control: ECS' CED is located in a private office, it is protected by doors closed with biometric access. Access to ECS' office is provided by biometric access.<br>• Hardware security: see above.<br>• Avoiding risk sources: all ECS' servers are isolated from the Internet to prevent risk sources not contemplated by other policies.<br>• Protection against not human risk sources: all ECS' servers are located in low seismic risk areas and physical servers are located in rooms with fire alarms.<br>• Security policies: ECS will be certified with ISO 27001 by the end of 2020.<br>• Privacy risks management: ECS does not collect users' personal data in the context of this project.<br>• Workstation management: Access to devices is protected with credentials or biometric data. Devices' disks are encrypted. Devices can be geolocated and reset remotely.<br>• Vulnerability: All ECS devices are configured to automatically fetch and install the latest available security patches.<br>• Protection against malware: all devices install anti-malware software.<br>• Security of communications: Data is transferred using secure and encrypted channels only. |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | Yes, security risk assessments target all possible impacts. |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes, all these aspects are considered during risk assessments. |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | Yes, this is a strict requirement for ISO 27001 standard certification. |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | Yes, ECS has defined the roles, tasks and responsibilities of personnel conducting risk assessments. |
| | | What could be the main impacts on the data subjects if the risk were to occur? | Their corporate email address may leak outside of EasyPIMS. |

| | | | |
|---|---|---|---|
| RISKS | Illegitimate access to data | What are the main threats that could lead to the risk? | Third parties trying to access and retrieve the data with malicious techniques. |
| | | What are the risk sources? | Malicious hackers or malicious members of the industry |
| | | Which of the identified planned controls contribute to addressing the risk? | Network security, access control (both physical and logical). |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible |
| | | Planned or Existing Mitigation measures to the foreseen risks? | In case a data breach occurs, ECS will immediately investigate the nature of the cause and contact all involved users to make them aware of the breach. |
| | Unwanted modification of data | What could be the main impacts on the data subjects if the risk were to occur? | Data subjects (data buyers) would see their privacy metrics corrupted. |
| | | What are the main threats that could lead to the risk? | Third parties interested in corrupting the data, attacks from hackers, ECS employee unwillingly modifying data. |
| | | What are the risk sources? | Malicious activity, human error. |
| | | Which of the identified planned controls contribute to addressing the risk? | Traceability of data access and backup of datasets may help to identify this type of modifications. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible |
| | | Planned or Existing Mitigation measures to the foreseen risks? | In case a data modification occurs, ECS can use periodic backups to recover modified data. |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | Data subjects (data buyers) would see their privacy metrics disappeared. |
| | | What are the main threats that could lead to the risk? | Third parties interested in deleting the data, attacks from hackers, ECS employee deletes data by mistake |
| | | What are the risk sources? | Malicious activity, human error. |
| | | Which of the identified planned controls contribute to addressing the risk? | Backup of datasets will allow us to make most of the data available almost immediately after an incident. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible |
| | | Planned or Existing Mitigation measures to the foreseen risks? | In case data disappears, ECS will immediately investigate the nature of the incident. |

## 5. IAB Spain

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | IAB Spain |
| CONTEXT | OVERVIEW | What is the processing under consideration? | Organization of events/workshops in which users, usually belonging to companies interested in the project, register and, where applicable, give specific consent to participate in future project activities |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | IAB SPAIN is configured as data processor. IAB SPAIN internal contact is Paula Ortiz and Miguel Herranz, legal representatives of IAB SPAIN |
| | | Are there standards applicable to the processing? | There are no standards to apply, but general best practices are considered: data is maintained encrypted and transferred using secure channels. Data is accessible and processed by authorised personnel only, possibly using automatic procedures |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | The data is collected through forms on the website. We are collecting, name, emails and the company to which the individual belongs. |
| | | How does the life cycle of data and processes work? | The data will be collected through IAB SPAIN website, encrypted and secured, it will be blocked till we need it for fulfilling the project goals. The Information will be deleted permanently at the finalisation of the project (approx. 2/3 years). |
| | | What are the data supporting assets? | The data will be collected through IAB SPAIN website<br>With secure HTTPS and firewall protocols. At first it will be stored at our Wordpress version of the web (IAB Spain manages the servers). The data will be removed from the wordpress version of IAB SPAIN website and download for encryption and for the local storage of the same or by hosting the encrypted information in a cloud container (owncloud: https://owncloud.com/ ) that have their servers in the EU. The data will be processed only by IAB SPAIN staff participating in the project. All servers have access control through users and passwords and also with physical access controls.<br>(Backups are performed from Monday to Thursday incremental backups at 22:00 and full backups on Fridays at 22:00.) |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | We will process data to manage the registration for the events If the user authorizes it, we will also process their data in order to informing and inviting them to future events organised by the consortium of PIMCity project in the framework of this project |
| | | What is the storage duration of the data? | The Information will be deleted permanently at the finalisation of the project (aprox 2/3 years). |
| | | What are the legal basis making the processing lawful? | Those established by the GDPR. Each form informs you of the purpose for which your data is collected, how to exercise your rights and finally asks for the **consent** of each user. |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | Yes, we collect the data meant for the goal of the project. Also, data minimisation is fulfil as the data we collect is minimum (mail, name and company) |
| | Fairness | Are the data accurate and kept up to date? | The data is provided by the user. If the user wants to correct their data they will be able to do so at any time as they are informed of it in compliance with the GDPR |

| | | | |
|---|---|---|---|
| | | In case you rely on consent, is it really free? | Yes, is really free. We do not have pre-checked consent boxes i.e:<br>☐Consiento el envío de comunicaciones relativas a actividades que se llevarán a cabo en el proyecto. Opt in check box for communicating further activities of the Project. |
| | | How do you document the that people gave it? | The checked boxes are registered via our servers. Only people who needs access to those consents can access the information. |
| | | How can they revoke their consent? | In all our communications there's a direct link for unsubscribe. That revocation is automatic. Also in our privacy policy we provide several direct communication channels to revoke consent (ordinary mail, email, phone…) |
| | | Could this generate chilling effects? | NO |
| | | Could this lead to discrimination? | NO |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Yes. In all of our workshops or activities we provide the user with several direct communication channels for this purpose |
| | Transparency | How will you tell people about your processing? | We have a minimum disclosure text with general information of the processing but we also provide more information. I.e.: "The data collected in this form will be processed by IAB Spain. We will process your data to manage your registration, remind you about the workshop and send you the information about it. If you authorize us to do so by checking the corresponding box, we will also process your data to invite you and inform you about the future events organised by the consortium of PIMCity project in the framework of this project (funded by the European Union Horizon 2020 Research and Innovation programme under the ICT theme: ICT-13-2018-2019, Grant Agreement No. 871370.) You can withdraw your consent and exercise your data protection rights by writing to us at iablegal@iabspain.es. For more information, please check the privacy policy of this workshop. privacy policy." |
| | | How do you make sure the information reaches the persons affected? | The information is always available for all users, via forms, via privacy policy… |
| | | Is the information you provide complete and easy to understand? | Yes. All the information that we provide to users is easy to understand and is provided in two languages (English and Spanish) |
| | | Is it (the provided information) targeted to the audience? | Yes. We have specific privacy policy and forms for all of the activities related to the project. |
| | | In case you defer informing people, how do you justify this? | In case we defer the information (not usual) we will follow the rules specified in the GDPR (One month as a general rule, but if we receive, for example, lots of petitions we can justify an answer more extended in time (one additional month with previous notification to the user) |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes, we have work with KU Leuven for identifying all purposes of our processing. Please check privacy policy. |
| | | Are all purposes compatible with the initial purpose? | Yes. Please check privacy policy. |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | No. We have separated databases for PIMCity project. |
| | | How can you ensure that data are only used for their defined purposes? | Only people who works on PIMCity project can access the data and all the people who works on the project knows for which |

| | | | |
|---|---|---|---|
| | | | purposes we need to treat the data. All IAB SPAIN employees has signed a data protection manual for this kind of purposes. |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | In general, the data that we collect will not be reused for other purposes not established in our privacy policy. If we need statistics we will only make available to the project general numbers (total of subscribers, total of attendees…) |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes, the data collected: name, email, company is enough and sufficient for the purpose of PIMCity project (Data minimisation) |
| | | Do the data you collect measure what you intend to measure? | Yes. |
| | | Are there data items you could remove without compromising the purpose of the process? | No. As the data we collect is minimum (mail, name and company) |
| | | Do you clearly distinguish between mandatory and optional items in forms? | Yes. Mandatory fields for registration are mail, name and company, optional items are presented via consent boxes (not prechecked) |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | If the case, we will delete all the data from our servers, from our security copies, and we will only keep general numbers (subscribers, attendees…) with any possibility of reintegration of the data. |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | As the data we collect is minimum (mail, name and company) it will be low impact. We are also clear in our communications regarding the scope of the same and how to unsubscribe/ delete the data |
| | | How do you ensure that the data you collect yourself are accurate? | The data are auto declared by the individuals |
| | | How do you ensure that data you obtain from third parties are accurate? | The data are auto declared by the individuals. No third parties only individuals declaring their data. |
| | | Do your tools allow updating / correcting data where necessary? | YES |
| | | Do your tools allow consistency checks? | YES |
| | Storage Limitation | Does EU legislation define storage periods for your process? | NO |
| | | How long do you need to keep which data? For which purpose(s)? | During the duration of the project. The Information will be deleted permanently at the finalisation of the project (approx. 2/3 years). |
| | | Can you distinguish storage periods for different parts of the data? | No, as the data we collect is minimum (mail, name and company) |
| | | If you cannot delete the data just yet, can you restrict access to it? | Yes, i.e.: if any individual have any complaint filed (DPA) we will block the data for access or modifications. |
| | | Will your tools allow automated erasure at the end of the storage period? | No. We will do it manually at the finalisation of the project. |

| | | | |
|---|---|---|---|
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | Yes. |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | Yes. Low in the context of the treatment of personal data that IAB Spain performs for PIMCity project |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes. |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | We have help from Tecnoderecho: https://sistemas.tecnoderecho.com/ (Data protection contracts in order) |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | Yes, we have legal profiles in IAB SPAIN that works within PIMCity project that could perform the risk assessment. We will be supported by tech companies. |
| RISKS | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | As the data that identify the users will be encrypted (SHA 256) there is low risk for users who provide their data (name email and company) |
| | | What are the main threats that could lead to the risk? | Third parties trying to access and retrieve the data with malicious techniques. Low risk as the data we collect is minimum (mail, name and company) |
| | | What are the risk sources? | Malicious hackers or malicious members of the industry. Low risk as the data we collect is minimum (mail, name and company) |
| | | Which of the identified planned controls contribute to addressing the risk? | All of them aims at avoiding this type of incident |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |
| | | Planned or Existing Mitigation measures to the foreseen risks? | All of them aims at this, activity log, access control, anti-hack tools, FW control |
| | Unwanted modificatio n of data | What could be the main impacts on the data subjects if the risk were to occur? | There is no risk to individual users as the DDBB will be encrypted and only managed by the relevant personnel of IAB SPAIN. In any case, impact is minimum as the data we collect is minimum (mail, name and company) |
| | | What are the main threats that could lead to the risk? | Not clear |

| | | | |
|---|---|---|---|
| | | What are the risk sources? | Malicious activity, human error |
| | | Which of the identified planned controls contribute to addressing the risk? | All of them aims at securing data storage and access. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited as the data we collect is minimum (mail, name and company) |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |
| | | Planned or Existing Mitigation measures to the foreseen risks? | All the measures previously exposed. Activity log, access control, anti-hack tools, FW control |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | There is low risk for individual users. Low impact on users as only their email, name and company will be collected. |
| | | What are the main threats that could lead to the risk? | Major Event<br>System bug<br>Third party unauthorised access<br>Human error |
| | | What are the risk sources? | Malicious activity, human error, Major Event |
| | | Which of the identified planned controls contribute to addressing the risk? | All of them aims at this, specially the backup plan and control access. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |
| | | Planned or Existing Mitigation measures to the foreseen risks? | All the measures the previously exposed. |

## 6. NEC

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | AUI |
| CONTEXT | OVERVIEW | What is the processing under consideration? | PIMCity has two main goals. The first one is about developing technical components that can help PIMS (Personal Information Management Systems) to be correctly implement. The second one is to develop and test a PIMS that will allow real users to trade their data with third parties, so, the data of private data is at the core of the project.<br>NEC will implement two main technical tasks in PIMCity involving data.<br>1) Privacy preserving data analytics<br>2) Knowledge extraction from network traces<br>For 1) the plan is to use publicly available datasets, but for 2) the plan is to install in potentially thousands of users a browser plugin that capture the hostnames visited by the users. That data will be then analysed to generate user profiles that will be traded with third parties. Of course, all with the consent of the user providing the data.<br>In the first phase of the project, the data will be stored at NEC premises (ANT testbed), only providing info the users through an interface, but in the final version (the whole PIMS implementation, with data being trade) the data will be stored, analysed and possible traded in a cloud provided by FastWeb. This cloud is to be defined in the second year of the project. |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | NEC will be the controller in the first part of the project, probably it will be joint controller with AUI, in charge of collecting the users. |
| | | Are there standards applicable to the processing? | We will follow the ISMS policies of NEC corporation. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | NEC will collect the hosts visited by different users with a browser plugin. That data will be used to generate user profiles. This kind of data can include sensitive data such as sexual orientation (i.e., if the user visits homosexual dating websites) if the user choses to upload this. |
| | | How does the life cycle of data and processes work? | The user installs a plugin in the browser and gives consent for the processing by doing this.<br>The data is periodically sent to the NEC premises using an encrypted connection.<br>The data is analyzed in NEC premises.<br>A profile of the user is returned to the user. |
| | | What are the data supporting assets? | The data will be stored in Debian machines. It will be stored in a database that could be MySQL or Mongodb |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | Yes, the users will give us consent for the collection and processing of the data. Before installing the plug-in, they are informed about the processing through the privacy policy. |
| | | What is the storage duration of the data? | As long as the account exists, up to a maximum of 3 years after the project (I.e. August 2025) |
| | | What are the legal basis making the processing lawful? | Consent |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | Yes, we will be collecting only the minimum data necessary to run the experiment, users are under no obligation to upload certain data. |
| | Fairness | Are the data accurate and | Yes, the user will have to possibility to update their data, and the |

| | | | |
|---|---|---|---|
| | | kept up to date? | browsing history is continuously updated |
| | | In case you rely on consent, is it really free? | Yes, users can choose if and which data to submit and/or upload. |
| | | How do you document the that people gave it? | They must execute an active action at the install (checkbox) |
| | | How can they revoke their consent? | By uninstalling the plug-in |
| | | Could this generate chilling effects? | Yes, but if people feel uncomfortable sharing their history, they can uninstall the plug-in |
| | | Could this lead to discrimination? | No |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Yes, that's the basis of the platform. |
| | Transparency | How will you tell people about your processing? | The users will receive all the information before the collection start, it will be included in the privacy policy of the plugin when they install it |
| | | How do you make sure the information reaches the persons affected? | Through the platform/mail |
| | | Is the information you provide complete and easy to understand? | Yes |
| | | Is it (the provided information) targeted to the audience? | Yes |
| | | In case you defer informing people, how do you justify this? | Not applicable. |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes |
| | | Are all purposes compatible with the initial purpose? | Yes |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | Yes, reidentification or combining the data with other data sets. |
| | | How can you ensure that data are only used for their defined purposes? | Usage of proper privacy preserving analytics, using the state of the art to mitigate the risk as much as possible. |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes |
| | | Do the data you collect measure what you intend to measure? | Yes |
| | | Are there data items you could remove without compromising the purpose of the process? | No. |
| | | Do you clearly distinguish between mandatory and optional items in forms? | Yes. |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re- | Applying the state of the art. |

| | | | |
|---|---|---|---|
| | | identification? | |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | / |
| | | How do you ensure that the data you collect yourself are accurate? | |
| | | How do you ensure that data you obtain from third parties are accurate? | |
| | | Do your tools allow updating / correcting data where necessary? | |
| | | Do your tools allow consistency checks? | |
| | Storage Limitation | Does EU legislation define storage periods for your process? | No |
| | | How long do you need to keep which data? For which purpose(s)? | During the duration of the project. The Information will be deleted permanently at the finalisation of the project (approx. 2/3 years). |
| | | Can you distinguish storage periods for different parts of the data? | No, as the data we collect is minimum (mail, name and company) |
| | | If you cannot delete the data just yet, can you restrict access to it? | Yes, i.e.: if any individual have any complaint filed (DPA) we will block the data for access or modifications. |
| | | Will your tools allow automated erasure at the end of the storage period? | No. We will do it manually at the finalisation of the project. |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | Yes. |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | Yes. Low in the context of the treatment of personal data that NEC performs for PIMCity project |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes. |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | Yes. |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | Yes, |
| | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | Potentially, sensitive personal information like the sexual orientation or health information could be disclosed. |
| | | What are the main threats that could lead to the risk? | Hackers breaking into the NEC premises systems |
| | | What are the risk sources? | Malicious actors |
| | | Which of the identified | Servers disconnected from the network, Servers only accessible to |

| | | | |
|---|---|---|---|
| RISKS | | planned controls contribute to addressing the risk? | authorized personnel, PII stored differently |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible |
| | | Planned or Existing Mitigation measures to the foreseen risks? | |
| | Unwanted modification of data | What could be the main impacts on the data subjects if the risk were to occur? | Wrong profiles received |
| | | What are the main threats that could lead to the risk? | Hackers breaking into the NEC premises system |
| | | What are the risk sources? | Hackers |
| | | Which of the identified planned controls contribute to addressing the risk? | Encrypted connection. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible. |
| | | Planned or Existing Mitigation measures to the foreseen risks? | |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | None |
| | | What are the main threats that could lead to the risk? | N/A |
| | | What are the risk sources? | N/A |
| | | Which of the identified planned controls contribute to addressing the risk? | Personal data stored in separate database. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible |
| | | Planned or Existing Mitigation measures to the foreseen risks? | |

## 7. TID

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | TID |
| CONTEXT | OVERVIEW | What is the processing under consideration? | Allow users to migrate their data to new platforms, in a privacy-preserving fashion. |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | The tool can optionally process it by filtering out (e.g., by applying differential privacy) sensitive information such as platform-inferred data (e.g., social interactions between users, or user unavailability due to event attendance) or user-inputted data (e.g., remove login credentials or debit card numbers), and outport it into a new Personal Information Management System (e.g., EasyPIMS) |
| | | Are there standards applicable to the processing? | No. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | The following data will be collected and processed:<br>- Social Media (e.g. Facebook, Twitter)<br>- Mobile Sensor Data (e.g. Motion Activity, Accelerometer, Gyroscope) - Banking<br>- Emails<br>- Calendar<br>All data will be temporarily saved until the user decides to empty the storage. There will be no expiration date.<br>Only the user that owns the data will have access. |
| | | How does the life cycle of data and processes work? | Users that want to migrate their data will first authenticate the data sources (e.g. Mobile phone, Email etc) and customize the data importation process (e.g. choose data fields that will be imported). When the user chooses to start the importation process, all data will be imported and saved internally in an encrypted database. Next, optionally, data transformation filters will be applied in order to anonymise and/or aggregate the data. Finally, the data will be exported (migrated) to the endpoint and all previously saved data will be erased. |
| | | What are the data supporting assets? | The data come from other Personal Information Management (PIM) systems (e.g. EasyPIMS, Mobile Phone, Facebook, Bank etc), and stored locally until the user decides to outport it to a new PIM platform and erase it. |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | The goal of the data portability tool is to realise the portability functionality of the PIMCity project, which is a requirement of GDPR. This is an explicit and legitimate goal in the context of the project execution. |
| | | What is the storage duration of the data? | Unlimited until the user decides to delete it. |
| | | What are the legal basis making the processing lawful? | A consent from the user will be acquired before the importation and processing process begins. |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | Yes. |
| | Fairness | Are the data accurate and kept up to date? | Data will remain accurate and up to date depending on the user's actions to repeat the migration process. |
| | | In case you rely on consent, is it really free? | Yes |
| | | How do you document the that people gave it? | There is no way to collect the user data without the user first gives their consent. |
| | | How can they revoke their consent? | They can request this from the relevant action in the UI. |
| | | Could this generate chilling effects? | No |

| | | | |
|---|---|---|---|
| | | Could this lead to discrimination? | No |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Yes, by leveraging the web interface, the data buyer will be provided the means to exercise their rights. |
| | Transparency | How will you tell people about your processing? | Information will be given in the UI. |
| | | How do you make sure the information reaches the persons affected? | The interface will include a ticking box to ensure the information has been read. |
| | | Is the information you provide complete and easy to understand? | Yes, the interface is undergoing proper tests to check presentation of information is quick and easy to understand. |
| | | Is it (the provided information) targeted to the audience? | Yes, only authorized roles corresponding to authenticated parties participating PIMCity will be able to access the information. |
| | | In case you defer informing people, how do you justify this? | N.A. |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes |
| | | Are all purposes compatible with the initial purpose? | Yes |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | No |
| | | How can you ensure that data are only used for their defined purposes? | Only the authorized people can access the data. |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | We do not plan to make data for these purposes at the moment. |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes |
| | | Do the data you collect measure what you intend to measure? | We do not measure anything |
| | | Are there data items you could remove without compromising the purpose of the process? | All data are needed |
| | | Do you clearly distinguish between mandatory and optional items in forms? | We do not use forms. All data are optional and up to the user to import. |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | We do not plan to keep such information. |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | The consequences are that the exported data will be not useful for the user. |
| | | How do you ensure that the data you collect yourself are accurate? | It is in the interest of the user to provide accurate data. |

| | | | |
|---|---|---|---|
| | | How do you ensure that data you obtain from third parties are accurate? | We do not obtain data from third parties. |
| | | Do your tools allow updating / correcting data where necessary? | Yes, data can be updated from new migrations. |
| | | Do your tools allow consistency checks? | No, we do not check for consistency. |
| | Storage Limitation | Does EU legislation define storage periods for your process? | DPS follows the general rule of storing the data for time needed to provide the service. |
| | | How long do you need to keep which data? For which purpose(s)? | Unlimited until the user decides to delete it. |
| | | Can you distinguish storage periods for different parts of the data? | No |
| | | If you cannot delete the data just yet, can you restrict access to it? | DPC can delete the data any time, but access to data is restricted by design. |
| | | Will your tools allow automated erasure at the end of the storage period? | No |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | No |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | N/A |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | N/A |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | N/A |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | No |
| RISKS | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | Exposure of personal data |
| | | What are the main threats that could lead to the risk? | Vulnerability in the system, Credential exposure |
| | | What are the risk sources? | Malicious hackers / attackers, User |
| | | Which of the identified planned controls contribute to addressing the risk? | Encryption, Anonymisation, access control management |
| | | How do you estimate the risk severity, especially according to potential | Important, It is an important risk as the system could possible store sensitive personal data. |

| | | | |
|---|---|---|---|
| | | impacts and planned controls? | |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited, The system will use state-of-the-art encryption technology and techniques. |
| | | Planned or Existing Mitigation measures to the foreseen risks? | |
| | Unwanted modification of data | What could be the main impacts on the data subjects if the risk were to occur? | Their data not being accurate. |
| | | What are the main threats that could lead to the risk? | Vulnerability in the system, Credential exposure |
| | | What are the risk sources? | Malicious hackers / attackers, User |
| | | Which of the identified planned controls contribute to addressing the risk? | Encryption, Anonymisation, access control management |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible, The system will use state-of-the-art encryption technology and techniques. |
| | | Planned or Existing Mitigation measures to the foreseen risks? | |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | Data lost. |
| | | What are the main threats that could lead to the risk? | Vulnerability in the system, Credential exposure |
| | | What are the risk sources? | Malicious hackers / attackers, User |
| | | Which of the identified planned controls contribute to addressing the risk? | Encryption, Anonymisation, access control management |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Negligible |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Negligible, The system will use state-of-the-art encryption technology and techniques. |
| | | Planned or Existing Mitigation measures to the foreseen risks? | |

## 8. UC3M

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | UC3M |
| CONTEXT | OVERVIEW | What is the processing under consideration? | Aggregated data from advertising platforms and openRTB platforms to obtain a large-scale database of audience values from market-side perspective. |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | Not identified any responsibility further than the typical ones related to efficiency and security in the data collection and storage. Since we do not envision to collect any personal data, no further privacy guarantees are foreseen. |
| | | Are there standards applicable to the processing? | There are no standards to apply, but general best practices are considered: data is maintained encrypted and transferred using secure channels. Data is accessible and processed by authorised personnel only, possibly using automatic procedures. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | We will collect aggregated data from advertising platforms and OpenRTB platforms about the value of audiences |
| | | How does the life cycle of data and processes work? | We will collect the data from advertising platforms and OpenRTB and store it in a database. The data will be made available through an API to third parties in the context of the PIMICITY project. Since historical data is of value to understand the evolution of audiences' value we do not plan to remove/destroy data and keep all data collected and processed for transversal/historical analysis.<br>We recall it is all aggregated data. |
| | | What are the data supporting assets? | The data come from advertising platforms and OpenRTB platorms. It will be stored in a database locally, aggregated and made available to an API to third parties so they can assess the value of different audiences.<br>We recall all collected data is aggregated data. |
| FUNDAMENTAL PRINCIPLES | General | Are the processing purposes specified, explicit and legitimate? | The goal of the data collection and processing is provide third party estimation about the value of different audiences. This is an explicit and legitimate goal in the context of the project execution. |
| | | What is the storage duration of the data? | The plan is to keep the data without removing it so we can do transversal analysis across time |
| | | What are the legal basis making the processing lawful? | We plan to collect only aggregated data. There is no need to request consent to end-users since no personal individual data is not planned to be collected |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | Yes, we collect the data meant for the goal of the project. Again no personal data is planned to be collected or processed. |
| | Fairness | Are the data accurate and kept up to date? | The data is planned to be collected periodically so the value of different audiences is kept up to date |
| | | In case you rely on consent, is it really free? | Our tool does not rely on consent |
| | | How do you document the that people gave it? | Our tool does not requires consent |
| | | How can they revoke their consent? | Our tool does not requires consent |
| | | Could this generate chilling effects? | Our tool does not requires consent |
| | | Could this lead to discrimination? | Our tool does not requires consent |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Our tool does not requires consent |
| | Transparency | How will you tell people about your processing? | Our tool does not collect or process personal data. The tool collect aggregated information from marketing platforms |

| | | | |
|---|---|---|---|
| | | How do you make sure the information reaches the persons affected? | Not applicable to our tool |
| | | Is the information you provide complete and easy to understand? | We provide the value of audiences as reported by marketing platforms. We do not provide information for all audiences existing in these platforms but for a subset of them as requested by the third party companies using our tool. In this context the information is complete and easy to understand for these third parties. |
| | | Is it (the provided information) targeted to the audience? | The information is targeted to other modules in PIMCITY that require a value of audiences, specifically the Trading Engine |
| | | In case you defer informing people, how do you justify this? | The question does not apply to our tool |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes. Just one purpose: obtaining the value of audiences in the considered marketing platforms. |
| | | Are all purposes compatible with the initial purpose? | Yes. |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | No. |
| | | How can you ensure that data are only used for their defined purposes? | Because the retrieved data has only price information, it does not include any other information, so that it can only be used to estimate the value of the associated audience |
| | | In case you want to make available / re-use data for scientific research, statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | We do not process personal data. This question does not apply to our tool |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes. |
| | | Do the data you collect measure what you intend to measure? | Yes. We want to measure the value of audiences and we retrieve exactly this metric from marketing platforms. |
| | | Are there data items you could remove without compromising the purpose of the process? | Yes. We can remove the data regarding the value of one audience without affecting the data stored about the value of other audiences. |
| | | Do you clearly distinguish between mandatory and optional items in forms? | We do not use forms |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | We only collect and process aggregated data, there is no risk of re-identification. |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | This question does not apply to our tool since we do not collect personal data |
| | | How do you ensure that the data you collect yourself are accurate? | We collect data from third parties |
| | | How do you ensure that data you obtain from third parties are accurate? | We conduct experiments to validate that the audience value provided by marketing platforms is coherent, by querying several times for the same audience in short periods of time, so that the received price must be the same or very similar |

| | | | |
|---|---|---|---|
| | | Do your tools allow updating / correcting data where necessary? | Yes. We can override the value of an audience at any time. |
| | | Do your tools allow consistency checks? | We conduct these tests ourselves as explained in our answer above |
| | Storage Limitation | Does EU legislation define storage periods for your process? | No as far as we are aware since we collect aggregate data |
| | | How long do you need to keep which data? For which purpose(s)? | The plan is to keep the data without removing it so we can do transversal analysis across time |
| | | Can you distinguish storage periods for different parts of the data? | No |
| | | If you cannot delete the data just yet, can you restrict access to it? | Yes, the access to the data can be restricted to a group of people/third parties we select |
| | | Will your tools allow automated erasure at the end of the storage period? | No. We have not provision for this because we do not store personal data. |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | We do not store personal data. This question does not apply to our tool. |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | Not applicable to or tool |
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | Yes |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | The UC3M research team of PIMCITY is responsible for this |
| RISKS | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | There is no risk to individual users, since only aggregated data is collected and processed |
| | | What are the main threats that could lead to the risk? | Third parties trying to access and retrieve the data with malicious techniques. |
| | | What are the risk sources? | Malicious hackers or malicious members of the industry |
| | | Which of the identified planned controls contribute to addressing the risk? | All of them aims at avoiding this type of incident |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |

| | | | |
|---|---|---|---|
| | | Planned or Existing Mitigation measures to the foreseen risks? | The data will be stored in servers which are protected with start-of-the-art security techniques implemented by the UC3M IT & Security department including firewall, traffic monitoring, server access control through users and passwords, servers allocated in rooms with physical access control, backups, etc |
| | Unwanted modification of data | What could be the main impacts on the data subjects if the risk were to occur? | There is no risk to individual users, since only aggregated data is collected and processed |
| | | What are the main threats that could lead to the risk? | Third parties trying to access and modify the data with malicious techniques. |
| | | What are the risk sources? | Malicious hackers or malicious members of the industry |
| | | Which of the identified planned controls contribute to addressing the risk? | All of them aims at avoiding this type of incident |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |
| | | Planned or Existing Mitigation measures to the foreseen risks? | The data will be stored in servers which are protected with start-of-the-art security techniques implemented by the UC3M IT & Security department including firewall, traffic monitoring, server access control through users and passwords, servers allocated in rooms with physical access control, backups, etc |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | There is no risk to individual users, since only aggregated data is collected and processed |
| | | What are the main threats that could lead to the risk? | Third parties trying to access and retrieve the data with malicious techniques. |
| | | What are the risk sources? | Malicious hackers or malicious members of the industry |
| | | Which of the identified planned controls contribute to addressing the risk? | All of them aims at avoiding this type of incident |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Limited |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Limited |
| | | Planned or Existing Mitigation measures to the foreseen risks? | The data will be stored in servers which are protected with start-of-the-art security techniques implemented by the UC3M IT & Security department including firewall, traffic monitoring, server access control through users and passwords, servers allocated in rooms with physical access control, backups, etc |

## 9. Wibson

| SECTIONS | SUB-SECTIONS | QUESTIONS | ANSWERS |
|---|---|---|---|
| | | | Wibson |
| CONTEXT | OVERVIEW | What is the processing under consideration? | The Trading Engine (TE) stores information about the Data Buyer, Offer, and transactions (Data Sellers that accepted to sell their Data), but not the Data itself. The primary purpose of storing it is transparency: The Seller can see what offers he has sold, to whom, and for what purpose. On the other side, the Buyer can see how many sellers have sold their data for the published offers. The TE collaborates with the Personal Consent Manager to reach the Sellers from a segment specified on the Offer, with the Personal Data Safe to grant the Buyer access to the Data once the transactions succeed, and also with the Data Valuation tools to determine the best Offer price for both parties.<br><br>The Taxonomy Generator (TG) receives as input raw personal data from different data sources and structures the data producing a schema definition for each one. The TG stores the schema definitions and exposes them through an API to let other processing systems consume them. The TG collaborates with the Data Portability and Control system that notifies the TG for new data. Also, the TG notifies the Persona Data Safe about new schema definitions from incoming data input. |
| | | What are the responsibilities linked to the processing? *(data owner/data subjects/controller/processor)* | Wibson act as the controller and processor of the activity. For each Data Buyer wanting to buy Data, a joint-controllership for each one of them is defined and regulated by a contract. |
| | | Are there standards applicable to the processing? | There are no standards to apply, but all systems, the TE, CM (consent manager) and TG, use industry standards on every aspect for storing and transmitting personal and anonymous data. |
| | DATA PROCESS AND SUPP. ASS | What are the data processed? (Are the data processed personal data as defined by GDPR?) | The TE stores:<br>1. Offer information:<br>• Identifier of the Buyer<br>• Explanation of how is going to use the Data<br>• Type of Data the Buyer wants to buy<br>• Segmentation information: "Buy <X type of Data> from Sellers from <Y country>". <Y country> is the segmentation information<br>• Access limits: "The Buyer has access to the data until the T date" or "The Buyer can access N times."<br>• Price: The value the Seller is going to receive in exchange for the Data<br>2. Whether the Data Seller accepted to sell its data or not<br>3. Transaction status<br>The information is stored for the duration of the activity, and both parties can access it.<br><br>The TG consumes raw data from the Data Portability and Control tool to generate and store the data source schema definition. This schema contains field names and a descriptor, but not the Data itself. Every system that needs to process data from the Personal Data Safe needs to know the structure of the information, so it needs to ask the specific schema definition.<br><br>The PCM stores:<br>1. Consent information:<br>i. Identifier of the Data Seller<br>ii. Status of the consent (enabled or disabled)<br>iii. Timestamp of the last change<br>iv. Data being enabled to share with this consent |

| | | | |
|---|---|---|---|
| | | | v. Purpose enabled for sharing this data with this consent.<br><br>The TM stores:<br>1. Tasks information:<br>i. Identifier of the Data Seller<br>ii. Identifier of the task<br>iii. Status of the task for this Data Seller (done or not done)<br>iv. Timestamp of when it was done (if any)<br>v. Points earned for doing the task<br>2. Sweepstakes:<br>i. List of identifiers of Data Sellers that have qualified for each sweepstake. |
| | | How does the life cycle of data and processes work? | The TE:<br>1. The Data Buyer publishes an Offer on the TE and notifies all Data Sellers involved about the Offer.<br>2. A group of Data Sellers decides to sell the information to the Data Buyer, and the TE registers this decision.<br>3. The TE closes the transaction by rewarding the Seller and also generating the necessary means to grant the Buyer access to the Data on the Data Safe.<br>The PCM:<br>1. A list of available consents is shown to the user.<br>2. The user selects the consents to be enabled.<br>3. The PCM enables each selected consent.<br><br>The TG:<br>1. The Data Portability and Control system notifies the TG about a new raw data to process.<br>2. The TG consumes the raw Data, generates the schema, and persists this information.<br>3. Another system is willing to process X data from the Personal Data Safe system. It requests to the TG the schema definitions from X data. The TG responds with the specified schema definition. |
| | | What are the data supporting assets? | The TE uses a local database to store transactional data and exposes transaction history to Data Sellers and Data Buyers through an API. A background job processing tool with specific database processes all the transactions allowing the system to scale.<br><br>The PCM uses a local database to store the status of all the consents and serves them through an API.<br><br>The TM uses a local database to store the list of tasks with their corresponding status per user and serves them through an API. |
| **FUNDAMENTAL PRINCIPLES** | General | Are the processing purposes specified, explicit and legitimate? | Yes, the data is processed to give the users their data rights and ownership. Users give explicit and legitimate consent on the use of their data.<br>The TE stores Offer information, registers Sellers acceptance and tracks transactional data for the purpose of making the core of the TE work and to provide transparency for both sellers and buyers. Data Buyers explicitly define the purpose and use of the data on the published Offers on the TE. The purpose is legitimate as data sellers can see what data has been sold, to whom and for what purpose.<br><br>The PCM stores consent information for the purpose of enabling the user to participate in Data Offers created in the TE. This is specified at the same moment the user sees the available consents and enables/disables them. The purpose is legitimate as Data Sellers can see the full list of consents (with their status) and change them at any time. |

| | | | |
|---|---|---|---|
| | | | The TM stores tasks information for the purpose of rewarding the user with points after certain achievements are done. In the case the user reaches a certain threshold, the user participates automatically in different sweepstakes. The purpose is specified and made explicit within the platorm where the user sees the tasks done and not done yet, plus the sweestakes status. The purpose is legitimate as Data Sellers achieve tasks by carrying out manual actions within the platform and see how each task status and points balance get updated. |
| | | What is the storage duration of the data? | Wibson stores data only for the time needed to provide the services based on TE and or the length of the project |
| | | What are the legal basis making the processing lawful? | The TE transactions occur only when the user gives his explicit consent, but data being exchanged is not collected in the system.<br>The PCM transactions occur only when the user enables or disables each consent explicitly and manually. No other data is being collected.<br>The TM transactions occur only when the user completes a task within the platform. No other data is being collected. |
| | | Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')? | For the TE, the data collected from both parties during a transaction is the minimal data that can be collected in order to provide transparency to them.<br><br>For the PCM, the data collected from Data Sellers is minimised and limited to the consents available in the platform.<br><br>For the TM, the data collected from Data Sellers is minimised and limited to the list of tasks and sweepstakes available in the platform. |
| | Fairness | Are the data accurate and kept up to date? | The TE collects data every time a Data Buyer publishes an Offer, and also every time a Data Seller accepts to sell a piece of information on the TE. These two cases are the reason to say that data is kept up to date and accurate. The TG also updates the generated schema when there is a change in the data sources available.<br>The TE collects data every time a Data Buyer publishes an Offer, and also every time a Data Seller accepts to sell a piece of information on the TE. These two cases are the reason to say that data is kept up to date and accurate. The TG also updates the generated schema when there is a change in the data sources available.<br><br>The PCM collects data every time the user changes a consent through the platform. The data stored reflects the last selections done by the user.<br><br>The TM collects data every time a user completes a task in the platform. The data stored reflects the last actions carried out by the user. |
| | | In case you rely on consent, is it really free? | Data processed by Wibson is provided by the company and users itself. Consent is given free. |
| | | How do you document that people gave it? | The data buyer and data seller will be provided with an interface (ticking a box) to give consent to data processing. Also in the platform there will be a history of all consents given. |
| | | How can they revoke their consent? | The company/data buyer and data seller providing data will be able to revoke consent at any time by accessing the web interface and removing their data. |
| | | Could this generate chilling effects? | No. |

| | | | |
|---|---|---|---|
| | | Could this lead to discrimination? | No. |
| | | Is it easy for people to exercise their rights to access, rectification, etc.? | Yes, by leveraging the web interface. Also will be implemented Wibson privacy solution to manage DSR. |
| | Transparency | How will you tell people about your processing? | In the case of the TE, Data Sellers accept to sell personal information to Data Buyers for a specific purpose and in exchange for some value or reward. The TG receives data from the Portability and Control tool and they get informed by this tool. |
| | | How do you make sure the information reaches the persons affected? | The interface will include a ticking box to ensure the information has been read. |
| | | Is the information you provide complete and easy to understand? | Yes, the interface is undergoing proper tests to check presentation of information is quick and easy to understand. |
| | | Is it (the provided information) targeted to the audience? | Yes, only authorized roles corresponding to authenticated parties participating PIMCity will be able to access the information. |
| | | In case you defer informing people, how do you justify this? | Wibson operates to always provide information. Deferring may happen for technical reasons which will be communicated to parties involved in the processing. |
| | Purpose Limitation | Have you identified all purposes of your process? | Yes |
| | | Are all purposes compatible with the initial purpose? | Yes |
| | | Is there a risk that the data could be reused for other purposes (function creep)? | Yes, but Wibson has defined proper policies to avoid data can be reused for other purposes. |
| | | How can you ensure that data are only used for their defined purposes? | Only authorized roles corresponding to authenticated parties participating PIMCity can access the data, and information about accesses (time, id of the party and data resource are logged). |
| | | In case you want to make available / re-use data for scientific research, | We do not plan to make data for these purposes at the moment. |

| | | | |
|---|---|---|---|
| | | statistical or historical purposes, what safeguards do you apply to protect the individuals concerned? | |
| | Data Minimisation | Are the data of sufficient quality for the purpose? | Yes, they are. |
| | | Do the data you collect measure what you intend to measure? | Yes, they do, and they are limited to that specific purpose. |
| | | Are there data items you could remove without compromising the purpose of the process? | No, all data items are needed. |
| | | Do you clearly distinguish between mandatory and optional items in forms? | Yes, Wibson will define which data are mandatory to provide for the purpose in the web interface made available to data buyers and data sellers |
| | | In case you want to keep information for statistical purposes, how do you manage the risk of re-identification? | Wibson does not plan to keep information for statistical purposes. |
| | Accuracy | What could be the consequences for the persons affected of acting on inaccurate information in this process? | Information obtained by companies could be inaccurate for their purpose and it will lack of interest in the future for the user's data |
| | | How do you ensure that the data you collect yourself are accurate? | It is in the interest of the data buyer and data seller to give accurate data to obtain value for it. |
| | | How do you ensure that data you obtain from third parties are accurate? | Wibson does not collect personal data from third parties and does not plan to do that in the future. |
| | | Do your tools allow updating / correcting data where necessary? | Yes, the web interface will allow data buyers and sellers to correct and update information any time. |

| | | | |
|---|---|---|---|
| | | Do your tools allow consistency checks? | The specific nature of the data collected by Wibson does not require to perform consistency checks, but both users and buyers can make updated in case it's data is not accurate |
| | Storage Limitation | Does EU legislation define storage periods for your process? | Wibson follows the general rule of storing the data for time needed to provide the service |
| | | How long do you need to keep which data? For which purpose(s)? | Wibson stores data only for the time needed to provide the services based on Trading engine, consent manager and taxonomy generator |
| | | Can you distinguish storage periods for different parts of the data? | Yes, for the different players (data sellers, data buyers, etc) |
| | | If you cannot delete the data just yet, can you restrict access to it? | Wibson can delete the data any time, but access to data is restricted by design. |
| | | Will your tools allow automated erasure at the end of the storage period? | No, Wibson has not automated this part of the processing but will use Wibson privact platform to manage user's data. |
| | Security | Do you have a procedure to perform an identification, analysis and evaluation of the information security risks possibly affecting personal data and the IT systems supporting their processing? | Yes. Security risk assessment are conducted periodically as a certification requirement.<br><br>Encryption<br><br>Information transferred between Data Seller and TE, and also Data Buyer and TE is routed through secure channels. Storage systems also work with encryption algorithms<br><br>Anonymisation<br><br>None of the identifiers used for the data subjects is based on personal data. With an TE generated data subject identifier it is not possible to recognize who is the final user.<br><br>Partitioning data<br><br>It is on the plans, and requires discussion, that each data subject identifier is generated for each of the transactions, so that it is not possible to correlate any data subject looking at the transactions.<br><br>Logical access control<br><br>Data Sellers and Data Buyers both communicate with TE to exchange data for some value or reward. Both parties will need to be authenticated using a protocol that ensures the highest levels of security, but is yet to be defined.<br><br>Traceability (logging) |

| | | | |
|---|---|---|---|
| | | | Interaction of data subjects is going to be tracked via an access log, taking the necessary precautions of not linking personal data like public IPs. Access logs are stored for the duration of the action.<br><br>Minimising the amount of personal data<br><br>The amount of personal data stored is minimal and the required to get both systems to work.<br><br>Operating security<br><br>Assets supporting personal data are not publicly available and are part of a Virtual Private Network accessible only from the application servers. Personal with specific credentials can access the application servers, and they can do so through a bastion host that changes its public address from time to time.<br><br>Clamping down on malicious software<br><br>Data is not transferred to less secure networks.<br><br>Managing workstations<br><br>Access to workstations is protected with credentials or biometric data. Workstations' physical drives are encrypted.<br><br>Website security<br><br>Wibson follows ANSSI guidelines to secure websites<br><br>Backups<br><br>Information on the supporting assets is backed up once every day, and the amount of backups held is limited to 1 month.<br><br>Maintenance<br><br>PIMCity's infrastructure provider manages the physical maintenance of hardware.<br><br>Network security<br><br>Assets supporting personal data are inside a Virtual Private Network accessible only from the application servers. Personal with specific credentials can access the application servers, and they can do so through a bastion host that changes its public address from time to time. Software for detecting changes on files and unauthorized accesses is installed on every service running.<br><br>Physical access control<br><br>PIMCity infrastructure provider guarantees no physical access to the infrastructure. |
| | | Do you target the impact on people's fundamental rights, freedoms and interests and not only the risks to the organisation? | Yes |

| | | | |
|---|---|---|---|
| | | Do you take into consideration the nature, scope, context and purposes of processing when assessing the risks? | Yes |
| | | Do you manage your system vulnerabilities and threats for your data and systems? | Yes |
| | | Do you have resources and staff with assigned roles to perform the risk assessment? | Yes |
| RISKS | Illegitimate access to data | What could be the main impacts on the data subjects if the risk were to occur? | Company employee info and end users data could be exposed outside EasyPim. |
| | | What are the main threats that could lead to the risk? | Third parties trying to access and retrieve the data with malicious techniques. |
| | | What are the risk sources? | Malicious hackers or malicious members of the industry |
| | | Which of the identified planned controls contribute to addressing the risk? | Network security, access control (both physical and logical). |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Working on Risk management based in GDPR |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Working on Risk management based in GDPR |
| | | Planned or Existing Mitigation | In case a data breach occurs, Wibson will execute it risk management plan and immediately investigate the nature of the cause and contact all involved users to make them aware of the breach. |

| | | | |
|---|---|---|---|
| | | measures to the foreseen risks? | |
| | Unwanted modification of data | What could be the main impacts on the data subjects if the risk were to occur? | Data subject would see inconsistency on their transaction history (TE), profile of the consent manager and/or schema generated (TG) |
| | | What are the main threats that could lead to the risk? | Third parties interested in corrupting the data, attacks from hackers, Wibson employee unwillingly modifying data. |
| | | What are the risk sources? | Malicious activity, human error. |
| | | Which of the identified planned controls contribute to addressing the risk? | Traceability of data access and backup of datasets may help to identify this type of modifications. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Working on Risk management based in GDPR |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Working on Risk management based in GDPR |
| | | Planned or Existing Mitigation measures to the foreseen risks? | In case a data modification occurs, Wibson can use periodic backups to recover modified data. |
| | Data Disappearance | What could be the main impacts on the data subjects if the risk were to occur? | Data subject would see inconsistency on their transaction history (TE), profile of the consent manager and/or schema generated (TG) |
| | | What are the main threats that could lead to the risk? | Third parties interested in deleting the data, attacks from hackers, Wibson employee deletes data by mistake |
| | | What are the risk sources? | Malicious activity, human error. |
| | | Which of the identified planned controls contribute to addressing the risk? | Backup of datasets will allow us to make most of the data available almost immediately after an incident. |
| | | How do you estimate the risk severity, especially according to potential impacts and planned controls? | Working on Risk management based in GDPR |
| | | How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls? | Working on Risk management based in GDPR |
| | | Planned or Existing Mitigation measures to the foreseen risks? | In case data disappears, Wibson will immediately investigate the nature of the incident. |

## 10. IMDEA

### 1. Overview

#### 1.1. What is the processing under consideration?

For the Data Valuation from User Perspective (DVTUP):

- Calculate the value that data is bringing to a certain machine learning, targeting or prediction model.

For the Data Provenance (DP):

- Obtain a watermarked version of a dataset, which allows to prove the ownership and provenance of such dataset.

### 1.2. What are the responsibilities linked to the processing?

Not identified any responsibility further than efficiency and security in the data collection and storage. Since we do not envision collecting any personal data, no further privacy guarantees are foreseen.

Note that for the Data Provenance tool, we read data from the Personal Data Safe. The Data Trading Engine should provide the coordinates to fetch such dataset so we read, watermark and return watermarked data to the Data Trading Engine. The dataset is distributed to the Data Trading engine but we do not store such data nor the original one we received unwatermarked. Ideally, when data needs to be returned and publicly stored in a filesystem, we may use asymmetric standard encryption to return it in an encrypted manner and with a watermark, so that only data owners can retrieve such data. We may store the hashed filename of the content we watermark to ensure its integrity but never the plaintext or user identifiers.

### 1.3. Are there standards applicable to the processing?

DVTUP: No.
DP: asymmetric encryption at the very end in case of need as mentioned.
Data, Processes and Supporting Assets

### 1.4. What is the data processed?

We collect aggregated anonymized mobility data and watermark browsing data from the Internet to feed the data valuation algorithms and data provenance tool.

### 1.5. How does the life cycle of data and processes work?

DVTUP: We collect the data from existing sharing or open data sites in the Internet. The value of such data is calculated through the processes and algorithms designed by IMDEA. The results are made available through an API to third parties in the context of the PIMICITY project. We do not plan to remove/destroy data and keep all data collected and processed for transparency analysis.
We recall it is all aggregated or anonymized data openly available in the Internet.
DP: We do not collect any data. We use a static, anonymized, previously collected dataset we have from eyeWnder research of professor N. Laoutaris.

### 1.6. What are the data supporting assets?

The data comes from existing sharing or open data sites in the Internet or a previous static dataset. It is stored in a database locally, processed and made available to an API to third parties so they can assess the value of different audiences or watermark the urls we provide.
We recall it is all aggregated or anonymized data openly available in the Internet or from a previous research paper for the case of DP.

## 2. FUNDAMENTAL PRINCIPLES

### 2.1. Are the processing purposes specified, explicit and legitimate?

The goal of the data collection and processing is providing to a third-party estimation about the value of different data sources in a certain prediction task, and derive an algorithm that would be useful in a loosely defined context. This is an explicit and legitimate goal in the context of the project execution.
The goal of the watermarking tool follows a similar logic but without any data collection, it just uses a previous dataset.

### 2.2. What is the legal basis making the processing lawful?

We plan to collect only aggregated or anonymized data. We understand that there is no need to request consent from end-users since personal individual data is not planned to be collected. At least to provide the prototypes and testing of the algorithms we define in both tools.

### 2.3. Are the data collected adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')?

Yes, we collect the data meant for the goal of the project. Again, no personal data is planned to be collected or processed.

**2.4. Are the data accurate and kept up to date?**

The data is planned to be collected once. Accuracy is relevant at the time of selecting which datasets to work with.

**2.5. What is the storage duration of the data?**

During the duration of the project.

## 3. RISKS

**3.1. Planned or Existing Measures**

The data is stored in servers which are protected with start-of-the-art security techniques implemented by the IMDEA IT & Security department including firewall, traffic monitoring, server access control through users and passwords, etc. We use servers at Polito to store the final version of our data management tools for demonstrating how they work, which probably also includes their related databases if any.

**3.2. Risk: Illegitimate access to data**

**3.3. What could be the main impacts on the data subjects if the risk were to occur?**

There is no risk to individual users, since only aggregated data is collected and processed in both tools' databases.

**3.4. What are the main threats that could lead to the risk?**

Third parties trying to access and retrieve the data with malicious techniques.

**3.5. What are the risk sources?**

Malicious hackers or malicious members of the industry.

**3.6. Which of the identified planned controls contribute to addressing the risk?**

All of them aim at avoiding this type of incident. Server security at both IMDEA and Polito is considered as well as encrypted traffic among tools if any.

**3.7. How do you estimate the risk severity, especially according to potential impacts and planned controls?**

Limited, demo data exposed in the worst case and no PII related to real data.

**3.8. How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls?**

Limited, demo data exposed in the worst case and no PII.

**3.9. Risk: Unwanted modification of Data**

**3.10. What could be the main impacts on the data subjects if the risk were to occur?**

Malfunctioning of some tools if they are overflown with a badly formatted, yet accepted input by the tools. There is no risk to individual users, since only aggregated or anonymous data is collected.

**3.11. What are the main threats that could lead to the risk?**

Null, since the results of the research is made public through papers, and referencing the already public datasets used in case anybody wants to reproduce the analysis. It is not possible to update or delete data from our tools' OpenAPI as we do not even require that feature ourselves.

**3.12. What are the risk sources?**

Null, since the results of the research is made public through papers, and referencing the already public datasets used in case anybody wants to reproduce the analysis. A malicious user would have to hack the whole system of the OpenAPI or find a vulnerability in the code released on public repositories and then again, data is put into locally secured database or decentralized file systems with administrative or individual access only.

**3.13. Which of the identified planned controls contribute to addressing the risk?**

All of them aim to secure data storage and access.

**3.14. How do you estimate the risk severity, especially according to potential impacts and planned controls?**

Limited

**3.15. How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls?**

Limited

**4. Data Disappearance**

**4.1. What could be the main impacts on the data subjects if the risk were to occur?**

There is no risk to individual users, since only aggregated data is collected and processed.

**4.2. What are the main threats that could lead to the risk?**

Null, since the results of the research is made public through papers, and referencing the already public datasets used in case anybody wants to reproduce the analysis.

**4.3. What are the risk sources?**

Null, since the results of the research is made public through papers, and referencing the already public datasets used in case anybody wants to reproduce the analysis.

**4.4. Which of the identified planned controls contribute to addressing the risk?**

All of them aim at securing data storage and access.

**4.5. How do you estimate the risk severity, especially according to potential impacts and planned controls**

Limited

**4.6. How do you estimate the likelihood of the risk, especially in respect of threats, sources of risk and planned controls?**

Limited

## ANNEX B – GUIDELINES FOR CONSENT MANAGEMENT

The guidelines provide non-exhaustive recommendations for consent management, taking into account the requirements provided by the Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data[64] (further as the GDPR) and the recommendations provided by the European

[64] Regulation 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

Data Protection Board (further as the EDPB).[65] Template informed consent form is provided as Annex F and shall be adjusted (tailored) by the partners' on a case-by-case basis taking into account particular details. More information was also provided to the partners in D7.2.

Consent is a concept used both in the GDPR and the Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (further as the e-Privacy Directive)[66] and is one of the legal grounds for personal data processing that may be given for one or more specific purposes.[67]

First, it is important that consent shall always be obtained before the controller starts processing personal data for which consent is needed.[68] Second, Art. 4(11) and Rec. 32 of the GDPR provides that consent shall be given '**by a clear affirmative act**' establishing a '**freely given**', '**specific'**, '**informed'** and '**unambiguous'** indication of the data subject's agreement to the processing of personal data. Besides the Art. 4(11) and Rec. 32 the GDPR provides requirements in Rec. 33, 42 and 43 as to how the controller must act to comply with the main elements of the consent requirement.[69] Each and every requirement shall be fulfilled, taking into account the relevant case law of the Court of Justice of the European Union and the recommendations (guidelines) provided by the competent institutions such as the EDPB or, if relevant, certain national data protection authorities. Generally, consent can only be an appropriate lawful basis if a data subject is offered control and is offered a genuine choice with regard to accepting or declining the terms offered or declining them without detriment.[70]

**Freely given**

For the consent to be freely given the data subject shall be able to genuinely exercise its autonomy. As the EDPB puts it, the element 'free' implies 'implies real choice and control for data subjects.[71] In relation to this, certain recommendations shall be taken into account, including but not limited:

- consent cannot be bundled up as a non-negotiable part of terms and conditions, e.g. consent for marketing cannot be hidden in the general terms and conditions of the website;
- there cannot be any element of pressure or influence, e.g. data subject cannot be required to consent for marketing in order to use basic features of an app;
- data controllers shall be particularly careful while relying on consent if they act as public authorities or employers since it is considered that there is often a clear imbalance of power in such relationships, hence consent may be regarded as not freely given; it is recommended to consider alternative legal basis for processing personal data;

personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119, 4.5.2016, p. 1–88.

[65] EDPB, Guidelines 05/2020 on consent under Regulation 2016/679, Version 1.1, adopted on 4 May 2020, available at <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf>, accessed 14/05/2020.

[66] 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) [2002] OJ L 201, 31.7.2002, p. 37–47. The GDPR conditions for obtaining valid consent are applicable in situations falling within the scope of the e-Privacy Directive. EDPB, Guidelines 05/2020 on consent, p. 5.

[67] Art. 6(1) GDPR.

[68] EDPB, Guidelines 05/2020 on consent, p. 18.

[69] Ibid, p. 5.

[70] Ibid, p. 4.

[71] Ibid, p. 6.

- data controllers shall avoid 'bundling' consent with acceptance of terms or conditions, or 'tying' the provision of a contract or a service to a request for consent to process personal data that are not necessary for the performance of that contract or service.[72]

*Cookie walls*

The EDPB has clarified that access to services and functionalities 'must not be made conditional on the consent of a user to the storing of information, or gaining of access to information already stored, in the terminal equipment of a user (so called cookie walls).'[73] Hence, it is not allowed to, e.g. block the access to the website unless the data subject accepts cookies.

**Specific**

The consent shall be provided for specific, particularly clear purposes (for each purpose separately) and the information regarding the consent shall be provided separately from all of the other information. As EDPB puts it, data subjects shall give their consent with the understanding that they are in control and their data will only be processed for those specified purposes.[74] E.g., requesting consent for 'business purposes' would not be specific enough.

**Informed**

In essence, it is considered that the consent cannot be considered meaningful in case data subject is not informed about the relevant details properly. In relation to this, a number of details shall be provided. The EDPB is of the opinion that at least the following information is required for obtaining valid consent:

- data controller's (or multiple (joint) controllers') identity(-ies);
- the purpose of each of the processing operations for which consent is sought;
- what (type of) data will be collected and used;
- the existence of the right to withdraw consent;
- information about the use of the data for automated decision-making[75] where relevant, and
- the possible risks of data transfers due to absence of an adequacy decision and of appropriate safeguards.[76]

Data controller is not obliged to provide a list of particular data processors, however shall provide a list of recipients or categories of recipients.[77]

The information regarding consent shall be provided separately from all of the other information such as terms and conditions, in a very clear and plain language.[78] Information regarding consent can be provided in written or oral statements, as well as in audio or video messages[79].

**Unambiguous – form of consent**

[72] For more details see EDPB, Guidelines 05/2020 on consent, p. 6-11.
[73] EDPB, Guidelines 05/2020 on consent, p. 11.
[74] Ibid, p. 13.
[75] In accordance with Art. 22 (2)(c) of the GDPR.
[76] As described in Art. 46 of the GDPR. EDPB, Guidelines 05/2020 on consent, p. 14.
[77] Ibid.
[78] Ibid, p. 16.
[79] Ibid.

Data controller shall be able to demonstrate that it is obvious that data subject consented to the particular processing of personal data. While it is rather self-explanatory when it comes to the written agreements which require signature, there are some grey zones in the electronic environment, discussed in detail below.

Consent may be given not only as a written statement but also by electronic means. This could include 'ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes, opt-out constructions, inactivity as well as merely proceeding with a service do not constitute consent'.[80] In case consent shall be given following a request by electronic means, 'the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided.'[81] In other words, it is recommended to interrupt the user experience to ensure there is no ambiguity.

Nevertheless, the above is not to suggest that ticking the box is the only option. The users may be suggested to swipe or to perform other actions as long as sufficient information is provided and these actions would be sufficient to signify an agreement to a particular request. However, 'scrolling or swiping through a webpage or similar user activity will not under any circumstances satisfy the requirement of a clear and affirmative action: such actions may be difficult to distinguish from other activity or interaction by a user'.[82]

Besides, consent may be given as an oral statement. However, in all of the cases while choosing the particular form data controller shall take into account that it shall be able to demonstrate that consent was given (with regard to the latter see also the paragraph on *demonstration* below).

*Explicit consent*

Data controller shall take into account that *explicit* consent is required in certain cases such as for processing of special categories of data or for automated individual decision-making, including profiling.[83] Data controller shall consider requiring signed written statement, filling of an electronic form, receiving an email, receiving an uploaded scanned document carrying the signature or using electronic signature.[84] EDPB also suggests to consider two stage verification involving communication via email and SMS messages.[85]

**Demonstration and withdrawal of consent**

For consent to be valid the controller shall be able to demonstrate that consent was given as well as to provide subject with the opportunity to withdraw it since obtaining consent should be a reversible decision[86].

*Demonstration*

Data controller is free to choose its own way of demonstrating that consent was given, however it shall prove not only, e.g., the affirmative action, but also that all of the conditions were fulfilled, e.g., that consent was informed, freely given, etc. After the end of the processing activity, such proof should be kept no longer then strictly necessary for compliance with a legal obligation or for the establishment, exercise or defence of legal claims.[87]

[80] Rec. 32 GDPR; EDPB, Guidelines 05/2020 on consent, p. 17.
[81] Ibid.
[82] Ibid, p. 18.
[83] For more details see Guidelines 05/2020 on consent, p. 19.
[84] Ibid.
[85] For particular guidelines see EDPB, Guidelines 05/2020 on consent, p. 19.
[86] Ibid, p. 5.
[87] In accordance with Art. 17(3)(b) and (e) GDPR. For more details see EDPB, Guidelines 05/2020 on consent, p. 21.

*Withdrawal*

Data controller shall ensure data subject can withdraw the consent any time. It is not required that the consent would be withdrawn in the same way as it was given, however it is important that this procedure is not complicated and certainly not requiring more effort than for giving the consent. E.g., if the consent was given through one mouse click, the same amount of effort shall be sufficient to withdraw (e.g., requesting to send an e-mail, let alone a registered mail would be excessive). Besides, data subject shall be able to withdraw free of charge and shall face no negative effects with regard to services.[88]

**Regarding the purposes of data processing – granularity requirements**

Consent 'should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them'.[89]

In particular, in case personal data will be processed for multiple purposes, data subjects shall be free to choose with regard to each purpose separately. For example, it cannot be required to consent to processing for marketing and analytical purposes with one tick.

***

Finally, obtaining consent does not negate or in any way diminish the controller's obligations to observe the principles of processing enshrined in the GDPR, 'especially Article 5 of the GDPR with regard to fairness, necessity and proportionality, as well as data quality'.[90] Although there is no specific time limit with regard to how long the consent may last, the EDPB recommends data controllers to refresh it periodically.[91]

For more details, including practical examples and details on specific areas of concern in the GDPR such as children and scientific research see the guidelines of the EDPB.[92] For the additional recommendations regarding consent management on the project's website(s) please also see Annex B of the D7.2.

Data controller shall also ensure compliance with the relevant national legal requirements.

**ANNEX C – GUIDELINES FOR PRIVACY POLICIES**

The guidelines provide non-exhaustive recommendations for privacy policies, taking into account the requirements provided by the Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data[93] (further as the GDPR). Template privacy policy is provided as Annex G and shall be adjusted (tailored) by the partners' on a case-by-case basis taking into account particular details.

The GDPR provides an obligation for the data controller to take into account the 'nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons'[94], and to implement 'appropriate technical and organisational measures' to ensure and to

[88] For more details see EDPB, Guidelines 05/2020 on consent, p. 22.
[89] Rec. 32 GDPR.
[90] EDPB, Guidelines 05/2020 on consent, p. 4.
[91] Ibid, p. 22.
[92] EDPB, Guidelines 05/2020 on consent.
[93] Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.
[94] Art. 24(1) GDPR.

be able to demonstrate that processing is performed in accordance with the GDPR accordingly. Where proportionate in relation to processing activities, these measures shall include the implementation of appropriate data protection policies.[95] In relation to this, there is a number of circumstances where privacy policies shall be considered to be necessary.

**Content**

In principle, a privacy policy shall provide all of the essential details related to data processing such as the categories of personal data being processed, processing purposes, third parties that may get access to personal data, etc. While considering which particular information shall be included in a privacy policy, certain provisions of the GDPR shall be taken into account.[96]

In the light of the PIMCity project it is relevant that personal data will be collected from the data subject. In such a case privacy policy shall be accessible no later than at the time when personal data are obtained, and shall provide these details:

1. categories of personal data collected (to be processed);[97]
2. controller's (and its representative's, if applicable) **identity and contact details**;[98]
3. **contact details of the data protection officer**, if applicable;
4. **purposes** of personal data processing**;**
5. **legal basis(-es)** for personal data processing;
6. **legitimate interests** of the controller or a third party, in case processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party;
7. **recipients** or categories of recipients of personal data, if any;
8. where applicable, **the fact** that the controller intends to transfer personal data to a third country or international organisation **and the existence or absence of an adequacy decision** by the European Commission, **or in the case the transfers would be based on appropriate safeguards or binding corporate rules,[99] or on the second subparagraph of Art. 49(1) of the GDPR, reference to the appropriate or suitable safeguards** and the means by which to obtain a copy of them or where they have been made available;
9. **period** for which personal data will be stored;[100]
10. reference to the right to request access to and rectification or erasure of personal data or restriction of processing or to object to processing as well as the right to data portability;
11. in case processing is based on **consent**, **right to withdraw** it;
12. the right to lodge a complaint with a supervisory authority;
13. whether the provision of personal data is (i) a **statutory** or **contractual requirement**, **or** (ii) a requirement **necessary to enter into a contract**, as well as (iii) whether the data subject is **obliged** to provide personal data and (iv) of the possible **consequences of failure to provide** such data;
14. the existence of **automated decision-making**, including **profiling**, referred to in the Art. 22(1) and (4) of the GDPR **and**, at least in those cases, **meaningful information about the logic involved**,[101] as well as the **significance and the envisaged consequences of such processing** for the data subject.[102]

[95] Art. 24(1) and 24(2) GDPR.
[96] Including but not limited Art. 12-15 GDPR.
[97] E.g. names, surnames, IP addresses, etc.
[98] The requirements listed in lines 2-14 are stemming from Art. 13 GDPR.
[99] As referred to in Art. 46 or 47 GDPR.
[100] In case it is not possible, the criteria used to determine that period shall be provided.
[101] With regard to alternative approaches to providing such information see, e.g. M. Brkan and G. Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas', 11 (2020) *European Journal of Risk Regulation*, https://doi.org/10.1017/err.2020.10.
[102] Art. 13 GDPR.

In case the controller would like to process personal data for a purpose other than that for which the personal data were collected, the controller shall provide the data subject prior to that further processing with information on that other purpose and with any relevant information identified in the lines 9-14 above.[103]

Where and insofar as data subject already has the information listed above, it is not necessary to provide it additionally.[104]

In case personal data have not been obtained from the data subject, the controller shall provide the data subject with the information required by the Art. 14 of the GDPR.[105]

**Form**

Information provided in a privacy policy shall be concise, transparent, intelligible, easily accessible, written in clear and plain language, particularly if addressed to a child, and free of charge.[106] The information may be provided in combination with standardised icons in order to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing. Where the icons are presented electronically, they shall be machine-readable.[107]

For the additional recommendations regarding cookies management on the project's website(s) please also see Annex B of the D7.2.

***

Data controller shall also ensure compliance with the relevant national legal requirements.

## ANNEX D – GUIDELINES ON DATA PROCESSORS AND DATA PROCESSING AGREEMENTS

### Guidelines on data processors and data processing agreements

The guidelines provide requirements for data processors and non-exhaustive recommendations for data processing agreements, taking into account the provisions of the Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data[108] (further as the GDPR). The project partners' should apply these guidelines in consultation with their legal departments and taking into account the specific details of the particular data processing operations.

*Requirements for data processors*

A project partner who acts as a data controller shall only use only a processor which provides sufficient guarantees to implement appropriate technical and organisational measures in such a manner that processing will meet the requirements of the GDPR and ensure the protection of the rights of the data subject (Art. 28(1) GDPR). According to the European Data Protection Board, the controller is responsible for

[103] Art. 13(3) GDPR.
[104] Art. 13(5) GDPR.
[105] Art. 14 GDPR.
[106] Art. 12 GDPR.
[107] Art. 12(7) GDPR.
[108] Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.

assessing the sufficiency of the guarantees provided by the processor and should be able to prove that it has taken all of the elements provided in the GDPR into serious consideration (Guidelines 07/2020 on the concepts of controller and processor in the GDPR issued by the European Data Protection Board (further as the EDPB) on 7 July 2021 (further as the EDPB Guidelines), p. 31). The processor should be able demonstrate the sufficiency of its guarantees and it may require an exchange of relevant documentation. For example, according to the EDPB, it will often require providing *privacy policy, terms of service, record of processing activities, records management policy, information security policy, reports of external data protection audits, recognised international certifications, like ISO 27000 series* (EDPB Guidelines, p. 31). The controller should assess such a sufficiency on a case-by-case basis, taking into account the nature, scope, context and purposes of processing as well as the risks for the rights and freedoms of natural persons (EDPB Guidelines, p. 31). It is particularly advisable to take into account the expert knowledge (e.g. technical expertise), reliability and resources, besides, the reputation may also be a relevant factor (EDPB Guidelines, p. 31).

A processor cannot engage another processor without prior specific or general written authorisation of the controller. In case a general written authorisation would be issued, the processor shall inform the controller of any intended changes concerning the addition or replacement of other processors, thereby giving the controller the opportunity to object to such changes (Art. 28(2) GDPR).

*Requirements for data processing agreements*

Processing shall be governed by a written agreement which, according to Art. 28(3) GDPR, sets out, among other things:

(i) subject-matter of the processing (for example, storing data of the platform);
(ii) duration of the processing (for example, the specific duration of the project);
(iii) nature of the processing (for example, the very specific operations performed in such a way that would allow third parties to understand both the content and the risks of processing);
(iv) purposes of the processing (for example, securely storing the data of the data subjects);
(v) type of personal data (for example, email address);
(vi) categories of data subjects (for example, platform users);
(vii) the obligations and rights of the controller (as provided in the GDPR; for guidance also see the EDPB Guidelines p. 35).

The processor shall not process the data except on instructions from the controller, unless required to do so by EU or Member State law (Art. 29 GDPR). It means that the controller should provide its processor with detailed instructions for processing activities (for guidance see the EDPB Guidelines p. 35-36.

According to Art. 28(3) GDPR, the agreement should provide, among other things, that the processor:

*(a) processes the personal data only on documented instructions from the controller, including with regard to transfers of personal data to a third country or an international organisation, unless required to do so by EU or Member State law to which the processor is subject; in such a case, the processor shall inform the controller of that legal requirement before processing, unless that law prohibits such information on important grounds of public interest;*
*(b) ensures that persons authorised to process the personal data have committed themselves to confidentiality or are under an appropriate statutory obligation of confidentiality;*
*(c) takes all measures required pursuant to Art. 32 GDPR;*
*(d) respects the conditions for engaging another processor under GDPR;*
*(e) taking into account the nature of the processing, assists the controller by appropriate technical and organisational measures, insofar as this is possible, for the fulfilment of the controller's obligation to respond to requests for exercising the data subject's rights laid down in Chapter III of the GDPR;*
*(f) assists the controller in ensuring compliance with the obligations pursuant to Art. 32 to 36 GDPR taking into account the nature of processing and the information available to the processor;*
*(g) at the choice of the controller, deletes or returns all the personal data to the controller after the end of the provision of services relating to processing, and deletes existing copies unless Union or Member State law requires storage of the personal data;*

(h) *makes available to the controller all information necessary to demonstrate compliance with the obligations laid down in this Article and allow for and contribute to audits, including inspections, conducted by the controller or another auditor mandated by the controller.*

With regard to the latter (point h), the processor shall immediately inform the controller if, in its opinion, an instruction infringes the GDPR or other EU or Member State data protection provisions.

Nevertheless, it should be noted that an agreement should elaborate in detail on each of the important elements of such agreements. As the EDPB provides very clearly, *the processing agreement should not merely restate the provisions of the GDPR: rather, it should include more specific, concrete information as to how the requirements will be met and which level of security is required for the personal data processing that is the object of the processing agreement. Far from being a pro-forma exercise, the negotiation and stipulation of the contract are a chance to specify details regarding the processing* (EDPB Guidelines, p. 34). Besides, taking into account the specific situation, additional details may have to be included.

According to the EDPB Guidelines, a data processing agreement shall be *in writing, including in electronic form, and be binding. The controller and the processor may choose to negotiate their own contract including all the compulsory elements or to rely, in whole or in part, on standard contractual clauses* (EDPB Guidelines, p. 4). Non-written agreements cannot be considered sufficient (EDPB Guidelines, p. 31). Although this particular agreement can be a part of a broader agreement, it is advisable to have all the relevant conditions concerning data processing in one place such as a separate chapter or an annex.

For detailed guidance also see the EDPB Guidelines p. 30-43.

**ANNEX E – GUIDELINES FOR JOINT CONTROLLERSHIP AGREEMENTS**

**Guidelines for joint controllership agreements**

The guidelines outline the circumstances in which organisations are considered to be joint controllers and provides non-exhaustive recommendations for agreements between/among them, taking into account the provisions of the Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data[109] (further as the GDPR). The project partners' should apply these guidelines in consultation with their legal departments and taking into account the specific details of the particular operations.

*Regarding the status of a joint controller*

The organisations (two or more) are considered to be joint controllers in case they *jointly determine* the purposes and means of processing. According to the Guidelines 07/2020 on the concepts of controller and processor in the GDPR issued by the European Data Protection Board (further as the EDPB) on 7 July 2021 (further as the EDPB Guidelines), joint participation *needs to include the determination of purposes on the one hand and the determination of means on the other hand*, besides, an important criterion is that the processing would not be possible without both parties' participation in the sense that the processing by each party is inseparable, i.e. inextricably linked (EDPB Guidelines, p. 3, 19). Processing for the same purposes essentially means that organisations involved process the data for the same or common purposes, or where the purposes are closely linked/complementary (EDPB Guidelines, p. 20). In the context of the platform developed within PIMCity it is particularly notable that the use of an already existing technical system such as a platform does not exclude joint controllership when users of the system can decide on the processing of personal data to be performed in this context (for more details see EDPB Guidelines, p. 21). For example, the

[109] Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.

Court of Justice of the European Union decided in Wirtschaftsakademie case (5 June 2018, Case C-210/16, ECLI:EU:C:2018:388) that the administrator of a fan page hosted on Facebook must be regarded as taking part in the determination of the means of the processing of personal data related to the visitors of its fan page, i.e. the administrator must be considered to be joint controller, because it was defining parameters based on its target audience and the objectives of managing and promoting its activities. As EDPB puts it, the *choice made by an entity to use for its own purposes a tool or other system developed by another entity, allowing the processing of personal data, will likely amount to a joint decision on the means of that processing by those entities* (EDPB Guidelines, p. 21). Nevertheless, the use of a common infrastructure such as a platform will not necessarily lead to joint controllership, especially if the processing is clearly separable and could be performed by one party without the other. The assessment of joint controllership should be carried out on a case by cases basis though, taking into account very particular factual circumstances.

In order to decide whether the organisations act as joint controllers it is important to assess, among other things: (i) whether they both/all have a decisive influence on how processing takes place, (ii) whether the processing would not be possible in the same way in case one of them would not participate (they are inseparably linked).

It is also notable that an entity will be considered as joint controller with the other(s) *only in respect of those operations for which it determines, jointly with others, the means and the purposes of the same data processing in particular in case of converging decisions*. What is particularly important in the context of the platform developed within PIMCity project is that in case *one of these entities decides alone the purposes and means of operations that precede or are subsequent in the chain of processing, this entity must be considered as the sole controller of this preceding or subsequent operation* (EDPB Guidelines, p. 20). *The fact that several actors are involved in the same processing does not mean that they are necessarily acting as joint controllers of such processing. Not all kind of partnerships, cooperation or collaboration imply qualification of joint controllers*, for example, *the exchange of the same data or set of data between two entities without jointly determined purposes or jointly determined means of processing should be considered as a transmission of data between separate controllers* (EDPB Guidelines, p. 24). Besides, joint controllership may also be excluded in a situation where several entities use a shared database or a common infrastructure, if each entity independently determines its own purposes (EDPB Guidelines, p. 24). There may be situations where various actors successively process the same personal data in a chain of operations, with independent purposes and independent means. In such a case the concerned organisations act as successive independent controllers. In relation to this, the organisation which is in charge of the platform should assess carefully to what extent and what kind of decisions is it making together with data buyers. Depending on the particular circumstances, different scenarios may not be excluded.

*Regarding the agreements between/among joint controllers*

In cases of joint controllership, the concerned organisations should determine their respective responsibilities for compliance with the obligations under the GDPR, particularly as regards the exercising of the rights of the data subject and their respective duties to provide the information referred to in Art. 13 and 14 GDPR. Put simply, it is advisable that the agreement would provide in detail how, when, by whom and what kind of information should be provided to the data subject. The arrangement may also designate a contact point for data subjects (Art. 26(1) GDPR), and the EDPB considers that it is advisable (EDPB Guidelines, p. 47). This arrangement shall also reflect the respective roles and relationships of the joint controllers vis-à-vis the data subjects, and the essence of such arrangement shall be made available to the data subject (Art. 26(2) GDPR). The GDPR does not specify what is considered to be *the essence of such arrangement,* however it is advisable to provide all the information referred to in Art. 13 and 14, and specifying which controller is responsible for which element. According to the EDPB Guidelines, p. 4, the distribution of responsibilities should cover other controller obligations such as *regarding the general data protection principles, legal basis, security measures, data breach notification obligation, data protection impact assessments, the use of processors, third country transfers and contacts with data subjects and supervisory authorities*. Yet, it should be noted that joint responsibility does not necessarily result into equal responsibility, and it may very well be that, depending on the degrees and scope of involvement, the levels of responsibility would be different (for further details see

EDPB Guidelines, p. 20). Although the form of joint controllership arrangements is not regulated by the GDPR, it is advisable to sign a binding written agreement. It should be noted, however, that the data subject may exercise his or her rights under the GDPR in respect of and against each of the controllers (Art. 26(3) GDPR).

More guidance on the particular provisions of the agreements may be found in the EDPB Guidelines, p. 43-48.

**ANNEX F – TEMPLATE INFORMED CONSENT FORM**[110]

**[Name of the Data Controller]**

**Consent for Processing of Personal Data**

This document provides you an opportunity to allow us to process your personal data and information that shall help you to decide whether you want to allow that.[111] In particular, this document provides details about (i) who would be controlling your data; (ii) for what purposes your data would be processed; (iii) what (type of) your data would be processed; and reminds that you always have a right to withdraw your consent.

**Your personal data would be processed by** [Name of the data controller], a [Company type], with registered addressed at [Address] (further as 'Controller' or 'We').

Your personal data, in particular **[Provide categories of personal data]** would be processed **for [Description of the purposes in detail].**

In case you would have any **questions** or **concerns** with regard to processing of your data, you can always contact the us by sending an e-mail [E-mail address of the data protection officer (team) or other responsible person]. You could also **withdraw your consent at any time** by sending an e-mail to the same e-mail address.

Detailed information about your rights and processing of your data could be found at [Link to a website, e.g. privacy policy].

In case you agree that we would process your data under the conditions described above, please kindly sign this document:

[110] [The template consent form shall be adjusted (tailored) by the partners' on a case by case basis taking into account particular details. This template consent form is drafted for the cases where there is no automated decision-making and no risks related to absence of an adequacy decision and of appropriate safeguards.]

[111] As required by the Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88, Art. 4(11), Rec. 32, and recommended by the European Data Protection Board, Guidelines 05/2020 on consent under Regulation 2016/679, Version 1.1, adopted on 4 May 2020, p. 14.

[In case there are several purposes, there shall be an opportunity to express will with regard to every purpose separately. In case consent would be requested by electronic means, data controller shall ensure all of the relevant information is provided, revealed in detail in Annex B].

_____

Name:
Date:

**ANNEX G – TEMPLATE PRIVACY POLICY**[112]

**[Name of the Data Controller]**

**Privacy Policy**

Last revised on [Date]

*Please read this privacy policy carefully so that you fully understand how we collect, use and store information about you. In case you accept this privacy policy, we assume that you understand and agree with all the details provided below. Occasionally, we may update this privacy policy, so please kindly visit our website to always find the latest version. We will contact you in case there will be important changes.*

**1. Introduction**

1.1. [Name of the data controller], a [Company type], with registered addressed at [Address] (further as 'the Controller' or 'we') protects your information and your privacy and processes your personal data following the requirements provided by the relevant laws.[113]

1.2. This privacy policy (further as 'the Privacy Policy') describes how we process your personal data. In particular, this Privacy Policy provides details on [Please review and adjust the section below carefully; please only leave those sections which are relevant, e.g. consider deleting sections 1.2.3, 1.2.6, 1.2.7, 1.2.8, 1.2.11 and/or 1.2.14 in case they appear irrelevant; remove unnecessary details such as phrases 'if applicable' and 'if any']:

1.2.1. personal data collected (to be processed);[114]
1.2.2. controller's (and its representative's, if applicable) **identity and contact details**;[115]
1.2.3. **contact details of the data protection officer**, if applicable;
1.2.4. **purposes** of personal data processing**;**

[112] [The template privacy policy shall be adjusted (tailored) by the partners' on a case by case basis taking into account particular details. This template privacy policy is drafted for the cases where personal data are collected from the data subject. It shall be accessible at the time when personal data are obtained.]
[113] As required by the Regulation 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) [2016] OJ L 119, 4.5.2016, p. 1–88.
[114] E.g. names, surnames, IP addresses, etc.
[115] The requirements listed in 1.2.2-1.2.14 are stemming from Art. 13 GDPR.

1.2.5. **legal basis(-es)** for personal data processing;
1.2.6. **legitimate interests** of the controller or a third party, in case processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party;
1.2.7. **recipients** or categories of recipients of personal data, if any;
1.2.8. where applicable, **the fact** that the controller intends to transfer personal data to a third country or international organisation **and the existence or absence of an adequacy decision** by the European Commission, **or in the case the transfers would be based on appropriate safeguards or binding corporate rules,**[116] **or on the second subparagraph of Art. 49(1)** of the Regulation 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 **(**further **as the GDPR), reference to the appropriate or suitable safeguards** and the means by which to obtain a copy of them or where they have been made available;
1.2.9. **period** for which personal data will be stored [in case it is not possible, provide the criteria used to determine that period shall be provided];
1.2.10. reference to the right to request access to and rectification or erasure of personal data or restriction of processing or to object to processing as well as the right to data portability;
1.2.11. in case processing is based on **consent**, **right to withdraw** it;
1.2.12. the right to lodge a complaint with a supervisory authority;
1.2.13. whether the provision of personal data is (i) a **statutory** or **contractual requirement**, **or** (ii) a requirement **necessary to enter into a contract**, as well as (iii) whether the data subject is **obliged** to provide personal data and (iv) of the possible **consequences of failure to provide** such data;
1.2.14. the existence of **automated decision-making**, including **profiling**, referred to in the Art. 22(1) and (4) of the GDPR **and**, at least in those cases, **meaningful information about the logic involved**,[117] as well as the **significance and the envisaged consequences of such processing** for the data subject.[118]

[To ensure that information is provided in an easily accessible manner, consider providing certain information in a scheme, e.g. in a table. For example, consider providing the information of section 2 and section 3 in a table of three columns and multiple lines, where the first column would provide inforamtion about the purposes of processing, the second – inforamtion about the particular data being processed for the respective purpose, the third – legal basis for processing of that particular data and reference to the GDPR].

**2. Which information do we collect?**

2.1. We collect and process [Please provide particular details with regard to personal data which is collected and will be processed, e.g. to inform data subject – name, surname, e-mail address; to handle queries – query, request or complaint, information, related to the query, request or complaint, communication with the Controller; to engage in legal proceedings relating to data subject – all of the information mentioned above, documents and attachments submitted by you, procedural documents and court documents; including but not limited].

**3. Why do we collect information about you?**

3.1. [Please provide clearly and separately in a form of list the particular purposes of personal data processing, e.g. performing data analyses (including anonymization and aggregation of personal data),

[116] As referred to in Art. 46 or 47 GDPR.
[117] With regard to alternative approaches to providing such information see, e.g. M. Brkan and G. Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas', 11 (2020) *European Journal of Risk Regulation*, https://doi.org/10.1017/err.2020.10.
[118] Art. 13 GDPR.

creating individual profiles and using and sharing the resulting data to third parties for commercialisation or research purposes; preventing illegal activities; complying with and enforcing any applicable laws].

3.2. [Please provide legal basis(-es) for personal data processing clearly and separately in a form of list. Please be kindly reminded that legal bases include (a) consent (with reference to Art. 6(1)(a) of the GDPR); (b) necessity for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract (with reference to Art. 6(1)(b) of the GDPR); (c) necessity for compliance with a legal obligation to which the controller is subject (with reference to Art. 6(1)(c) of the GDPR); (d) to protect the vital interests of the data subject or of another natural person (with reference to Art. 6(1)(d) of the GDPR) [unlikely to be relevant]; (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller (with reference to Art. 6(1)(e) of the GDPR); (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child; in the latter case (of legitimate interests, please clarify them in detail[119]] (with reference to Art. 6(1)(f) of the GDPR). Please note the latter may cover processing queries, requests and complaints submitted by data subject for the purposes of handling them, as well as processing of various documents submitted by data subject or concerning data subject for the purposes of engaging to legal proceedings, including but not limited.

**4. Which information do you have to provide and why?**

4.1. You should provide us with the information that is necessary to handle your complaints, queries and requests. In case you would not provide us with such information, we would not be able to handle your complaints, queries and requests.

**5. Is your information shared with anyone else?**

5.1. [Please identify the recipients or categories of recipients of personal data.]

5.2. We may also share your personal data to courts and other institutions or subjects when it is required by law.

5.3. Other than as set out in this Privacy Policy, we shall not disclose your personal data to any third parties without obtaining your prior explicit consent unless that would be required by law.

**6. Is your information transferred to a third country or international organisation?**

6.1. [Please provide information regarding personal data transfers to a third country(-ies) and/or international orgnisation(s), including information indicated in 1.2.8; alternatively remove this section].

**7. For how long do we keep your information?**

7.1. [Please provide information on how long personal data will be stored [in case it is not possible, provide the criteria used to determine that period shall be provided].

**8. How do we secure your information?**

8.1. [Please provide information about the security measures].

[119] Art. 6(1) GDPR.

8.2. The Controller shall take appropriate administrative, technical and organizational measures against unauthorized or unlawful processing of any personal data or its accidental loss, destruction or damage, access, disclosure or use.

8.3. In the event of and following discovery or notification of a breach of the security of the personal data, or access by an unauthorised person, the Controller shall notify you if the breach is likely to affect your privacy.

**9. How do we manage cookies? [Remove if not applicable]**

[To ensure that information is provided in an easily accessible manner, consider providing certain information in a scheme, e.g. in a table.]

9.1. [Please provide information about the cookies – (i) cookie name, (ii) cookie category, (iii) cookie purpose and (iv) cookie expiry, each respectively].

9.2. We will not use cookies for which your consent is necessary without your consent.

9.3. Please note that you can configure your browser to decline all or some cookies as well as to ask your permission for accepting them.

9.4. To understand and control cookies of the other companies or institutions (so-called third-party cookies), please read the policies of such third-parties.

**10. What are your rights?**

10.1. You can ask the Controller whether it processes the data about you and, if yes, you can request access to that data**.**
10.2. You can ask the Controller to correct the inaccurate data about you.
10.3. You can ask to Controller to delete the data about you and the data shall be deteled in case there are no legal basis for the Controller to process it.
10.4. You can ask the Controller to restrict processing of your data (i) in case you contest the accuracy of the data, (ii) in case the processing is unlawful and you oppose the erasure of it, (iii) in case the Controller no longer needs the personal data but they are required by the data subject for the legal claims, or (iv) in case you have objected to processing pursuant to Art. 21(1) of the GDPR pending the verification whether the legitimate grounds of the controller override those of the data subject.
10.5. You can object processing your data in case it is processed for third party's or Controller's interests.
10.6. You can ask to transfer (receive) your data which is processed automatically and is provided to us by your consent or under the contract.
10.7. You can withdraw your consent given to us regarding processing of your personal data.
10.8. You have the right to request us not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning you or similarly significantly affects you [if applicable].
10.9. To complain to a supervisory authority and to seek a judicial remedy.

[Name of the relevant supervisory authority]
[Address of the relevant supervisory authority]
[Contacts of the relevant supervisory authority]

**How can you contact us?**

**Data Controller**
[Name of the data controller], a [Company type], with registered addressed at [Address]

[Contact details]

**Data Protection Officer**, if applicable
[Name] [Contact details]

## ANNEX H – PARTNERS' INDIVIDUAL INPUT FOR THE DATA MANAGEMENT PLAN

### 1. DATA SUMMARY

### 1.1. Politecnico di Torino

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

The purpose is to devise and test algorithms to automatically define privacy tags. Privacy tags refers to website and webservices, and not to individuals. This is one of the specific goals defined in WP2.

*What types and formats of data doesthe project generate/collect?*

We collect information instrumental to check if and how the website or web service collects and exchanges eventual personal information. Privacy tags summarize the output of the algorithms.

*Do you re-use any existing data and how?*

We have historical web crawling archives that have been performed in the past. These contains a snapshot of web pages done through the time using automatic web crawlers. Politecnico di Torino acts as data controller and data processor for the EasyPIMS demonstrator and is the only partner that has access to the personal data collected by the platform.

*What is the origin of the data?*

Some have been collected by POLITO internally by automatic web crawlers running on clusters of computers. Other are coming from public repositories such as https://web.archive.org. We plan to continue and enrich the web-crawling archives. For the data collected by the EasyPIMS platform, the data is being uploaded by end users which have the right to cancel, modify and withdraw from the demonstator.

*What is the expected size of the data?*

The size of the web archives can be very large, up to several terabytes of data depending on the extensiveness of the collection, and the frequencies at which these are collected. The size of the privacy tags archive is much smaller, in the order of few gigabytes. For the EasyPIMS demonstrator, the total amount of data collected by the platform is in the order of some gigabyte in total.

*To whom might it be useful ('data utility')?*

These web archives are commonly used for running algorithms to detected privacy violations from website and webservices.
These are also useful to check website performance (e.g., web page load time), and possible issues with pages (e.g., errors for missing objects).

Privacy tags are useful for the project goals and for anyone interested in the web data economy.

Data from the EasyPIMS demonstrator will not be made accessible to third parties or other partners to avoid the leakage of personal information of end users.

### 1.2. NEC Laboratories Europe

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

NEC collects the hosts visited by users to profile their interests. It is one of the main tasks parts of T4.4 and key of the personal data avatar.

*What types and formats of data will the project generate/collect?*

The data is composed by sequences of host, it is stored in a MySQL database.

*Will you re-use any existing data and how?*

We re-use anonimyzed dataset containing traffic logs.

*What is the origin of the data?*

The users.

*What is the expected size of the data?*

In the order of Gb of data.

*To whom might it be useful ('data utility')?*

To us.

### 1.3. Ermes Cyber Security

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

ECS designs, develops and tests algorithms to automatically generate Privacy Metrics (D-PM) which ware ultimately used to generate Transparency Tags. Transparency Tags describe which personal data is collected by websites and web services (e.g., browsing history). As such no individual is involved and no personal data is collected in this task. This is one of the specific goals defined in WP2.

*What types and formats of data will the project generate/collect?*

ECS collects data generated by its fleet of automatic web scrapers, which is then processed by automatic algorithms to identify which personal data web services collect and how. The result of this analysis, the D-PM, is provided in JSON format, and stored and distributed using state-of-the-art database technologies.

*Will you re-use any existing data and how?*

ECS re-uses data collected in the past using its fleet of automatic web scrapers. This contains information about web sites including the code used to generate the page (e.g., HTML and Javascript) as well as logs describing the APIs executed by the browser to generate the page.

*What is the origin of the data?*

ECS has collected such data internally, using its fleet of automatic web crawlers running on clusters of computers deployed in the cloud.

*What is the expected size of the data?*

ECS collected few terabytes of data so far, but such size varies depending on a number of variables (e.g, the number of browsed sites, number of samples, iterations, etc.). The size of the resulting dataset containing D-PMs is expected to be much smaller (few GBs).

*To whom might it be useful ('data utility')?*

Apart from the generation of D-PMs, collected data is used by ECS to feed several algorithms with different purposes such as website classification and vulnerability assessment.

D-PMs are the fundamental brick at the base of Transparency Tags and is used by the partners to achieve the project goals, and by anyone interested in understanding the usage of data in the web.

### 1.4. IMDEA Networks

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

IMDEA requires data to test and demonstrate the functionality of data valuation tools for combinations of sources from the users' perspective and the proposed data marketplace concepts provided. The objective of such processing is to find the response of an AI/ML algorithm to different combinations of inputs for a set of available sources or users, to calculate the value they are bringing to the specific task.
IMDEA also develops methodologies to modify the traded dataset and add some kind of watermark to the dataset that eventually helps track potential leakages of information by trusted partners. Should a copy of that information be made public in the Internet, the watermark should help identify the entity that shared it.

*What types and formats of data will the project generate/collect?*

IMDEA leverages existing public data available on the Internet that may eventually resemble the kind of information that the PDK or the EasyPIMS would be managing and trading. This is in general structured data, for instance anonymized mobility data or CDRs (call detail records) showing the mobility of people within an area, such as a city. More details could be provided once the specific use cases to implement are clear. As of now, we are working with this mobility information which in all cases is totally anonymized.

*Will you re-use any existing data and how?*

All the tests are done by reusing public data, by downloading copies of the target datasets to local development environments and using them in the testbeds.

*What is the origin of the data?*

Public data available in search engines (e.g. Google datasets) or provided by public entities (e.g. https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page). Such datasets are made public without including identifiable personal data for developers and data scientist to test their algorithms.

*What is the expected size of the data?*

It will strongly depend on the use case, ranging from MB to tens of GB (e.g. mobility data).

*To whom might it be useful ('data utility')?*

Mobility data is useful for city planners or transportation enterprises at the time of planning resources or operations in their enterprises. CDR from mobile operators are being actively used in the billing process or for network planning purposes. Governments are using anonymized mobility information to track mobility in the city, for instance recently to measure the degree of mobility of the population during the Coronavirus outbreak.

### 1.5. Universidad Carlos III de Madrid

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

The purpose is to identify the value that market assigns to end-users' value. This is one of the specific goals defined in WP3.

*What types and formats of data will the project generate/collect?*

We collects data regarding the value/price of audiences (i.e., user profiles) in online advertising platforms.

*Will you re-use any existing data and how?*

We have some data collected before the start of the project, in the context of TYPES and SMOOTH EU projects and we might reuse it. It is of the same nature and type than the one we plan to collect in PIMCITY.

*What is the origin of the data?*

Advertising Platforms (Closed ones as well as OpenRTB based).

*What is the expected size of the data?*

Between ten and thousands of GBs of data.

*To whom might it be useful ('data utility')?*

To PIMs in order to handle the offer of data on behalf of their user base. Third parties willing to bid (or request the use) of data in exchange of a monetary transaction.

### 1.6. Telefónica Investigación y Desarrollo

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

We collect and process data to enable data migration to new platforms in a privacy-preserving fashion. More specifically users are able to download their data from other Personal Information Management (PIM) systems (e.g. Mobile Phone, Email etc.), optionally filter out sensitive information, and outport into a new PIM system.

*What types and formats of data will the project generate/collect?*

It depends on the format that the 3rd party systems provide their data. In most cases this are textual type with formats such as JavaScript Object Notation (JSON) and Comma Separated Values (CSV).

*Will you re-use any existing data and how?*

No.

*What is the origin of the data?*

3rd party Personal Information Management (PIM) systems (e.g. Facebook, Mobile Phone, Email etc.)

*What is the expected size of the data?*

It tends to be hundreds of megabytes per user.

*To whom might it be useful ('data utility')?*

This tool benefits the users wishing to migrate their data to other platforms.

### 1.7. Fastweb

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

Fastweb collects or processes data in two ways:
Process data uploaded to PIMCity modules developed by other parties in the Consortium, because Fastweb hosts some PIMCity modules on its cloud computing infrastructure. This has the purpose of testing the modules' functionality and enable interaction between modules and, potentially, third parties.
Collect aggregated network data like the number of users or traffic share of select websites or services.

*What types and formats of data will the project generate/collect?*

The type and format of data that is processed in PIMCity modules running on Fastweb's cloud computing infrastructure using a custom data structure specifically designed in the project. The data is made accessible, previous authentication, via APIs as documented in the PDK documentation.

Fastweb collects aggregate network traffic data (e.g. traffic share of a particular website or service, bandwidth usage over time, number of users over time). The aggregation occurs in the dimension of users, i.e.: we do not collect or share any individual user identifier, but only user counts or session counts or bandwidth per website/service, possibly over time.

*Will you re-use any existing data and how?*

Fastweb might re-use some of the aggregated network traffic data it has been collecting for network capacity planning purposes, if useful for processing in PIMCity modules.

*What is the origin of the data?*

The data processed in PIMCity modules hosted in Fastweb's cloud computing infrastructure comes from the sources of said modules, developed and controlled by other Consortium parties. The data collected directly by Fastweb comes from network sensors installed in Fastweb's customer network.

*What is the expected size of the data?*

After some initial testing, we expect the data size to not exceed 10GB in total.

*To whom might it be useful ('data utility')?*

The data is useful to the developers of PIMCity modules to enrich the data in their modules.

### 1.8. LSTech ESPANA

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

LSTech provides applications/ mechanisms for data aggregation and anonymization and defines metadata schema for allowing the importing of personal data to the PIMCITY platform. LSTech does not need to collect or to have access to real personal data during the project. Test/ synthetic/ fake data can be used to test the applications.

LSTech also builds user interfaces in the form of dashboards to allow the visualizations of the usage of personal data. If needed, LST has access to the data (that is collected and stored by other partners) in order to verify and test the efficiency and validity of these interfaces.

*What types and formats of data will the project generate/collect?*

LSTECH uses data collected and used by IMDEA since both are participating in the same task. The types and formats of data that is collected/ used by IMDEA are mentioned in the respective section (1.2.4).

*Will you re-use any existing data and how?*

LSTECH uses data collected and used by IMDEA since both are participating in the same task.

All the tests are done by reusing public data, by downloading copies of the target datasets to local development environments and using them in the testbeds.

*What is the origin of the data?*

LSTECH uses data collected and used by IMDEA since both are participating in the same task.

Public data available in search engines (e.g. Google datasets) or provided by public entities (e.g. https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page). Such datasets are made public without including identifiable personal data for developers and data scientist to test their algorithms.

*What is the expected size of the data?*

It will strongly depend on the use case, ranging from MB to tens of GB (e.g. mobility data).

*To whom might it be useful ('data utility')?*

For the task T4.4, mobility data is used by LSTECH and IMDEA.

Mobility data is useful for city planners or transportation enterprises at the time of planning resources or operations in their enterprises. CDR from mobile operators are being actively used in the billing process or for network planning purposes. Governments are using anonymized mobility information to track mobility in the city, for instance recently to measure the degree of mobility of the population during the Coronavirus outbreak.

### 1.9. KU Leuven – CiTiP

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

KU Leuven – CiTiP provides ethical and legal guidance to the other PIMCity partners. It does not collect personal data, does not collect or generate datasets, does not create algorithms, computer software and the like. KU Leuven – CiTiP collects and generates research data, however, which shall contribute to the fulfilment of the overall objectives of the Project.

*What types and formats of data will the project generate/collect?*

KU Leuven – CiTiP generate research data in the form of deliverables which is saved primarily in .docx and .pdf formats.

*Will you re-use any existing data and how?*

KU Leuven – CiTiP re-uses the generated research data for further research.

*What is the origin of the data?*

KU Leuven – CiTiP is generating research data relying on its own expertise in the subject matter.

*What is the expected size of the data?*

To be defined at the final stages of the project according to the documentation produced

*To whom might it be useful ('data utility')?*

Primarily for subjects researching privacy, data protection and intellectual property law.

### 1.10. Asociación de Usuarios de Internet

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

The purpose of the data collected by AUI is to engage users to participate in the project by filling out surveys, participating in focusgroups and to participate as betatesters.

*What types and formats of data will the project generate/collect?*

Three types of data are collected:

- Identifying data (name, email, phone) collected exclusively to contact participants who wish to give them voluntarily

- Sociodemographic data (age, gender, country, marital status, professional status, level of education, country, state, city, language) for the elaboration of studies and evaluation of the use of the different tools.

- Use and activity data (log files)

As for the formats these are collected in records and tables of mysql type databases in which the identification data are always stored encrypted.

*Will you re-use any existing data and how?*

No.

*What is the origin of the data?*

The data are provided by the users themselves when filling in the forms or registering on the website.

*What is the expected size of the data?*

The number of users will be less than 3000 and therefore the associated data can be less than 5 Megabytes.

*To whom might it be useful ('data utility')?*

For the set of partners that can through this data understand how different types of users perceive and iteract with the tools developed for them in the project.

### 1.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*What is the purpose of the data collection/generation and its relation to the objectives of the project?*

Contact information for workshops.

*What types and formats of data will the project generate/collect?*

Contact information.

*Will you re-use any existing data and how?*

No

*What is the origin of the data?*

Participants of workshop.

*What is the expected size of the data?*

Little.

*To whom might it be useful ('data utility')?*

Users and us to contact users.

### 1.12. WIBSON

Wibson does not process any data related to the PIMCity project.

### 1.13. TAPTAP

TAPTAP does not process any data related to the PIMCity project.

## 2. FAIR DATA

### a. Making data findable, including provisions for metadata

#### 2.1.1. Politecnico di Torino

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

No. The data we plan to collect, process and produce is aggregated data.

*What naming conventions do you follow?*

Given neither a standard format nor a best-practice is available, we use custom and well-documented format for naming datasets.

*Will search keywords be provided that optimize possibilities for re-use?*

We implement a searchable database to access and retrieve the privacy tags. The access to this data is available through open and standard API offered by Ermes and documented in the PDK.

*Do you provide clear version numbers?*

Yes, we identify the different versions of the privacy tag API with a version number.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

No standard metadata exists for the privacy tags.
Similarly for Web archives, no standard metadata exists.

### 2.1.2. NEC Laboratories Europe

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

No.

*What naming conventions do you follow?*

It is stored in a MySQL database.

*Will search keywords be provided that optimize possibilities for re-use?*

No, we will not ask the user for consent to reuse.

*Do you provide clear version numbers?*

A single version is generated.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

No.

### 2.1.3. Ermes Cyber Security

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

No. They are not.

*What naming conventions do you follow?*

It has not been defined yet.

*Will search keywords be provided that optimize possibilities for re-use?*

Privacy Metrics is stored in a searchable database and made available through APIs.

*Do you provide clear version numbers?*

Yes, both Privacy Metrics and APIs are versioned.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

Metadata for Privacy Metrics have not been defined yet.

### 2.1.4. IMDEA Networks

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

The objective of the modules produced by IMDEA is to be used internally by other modules of the project. For instance, the relative values calculated for the different datasets contributing to a piece of traded data might be used by the trading engine in order to share the reward with data providers.
Regarding the watermarking, it just includes some tracking information in the data, but it does not alter the original metadata and/or its discoverability properties.

*What naming conventions do you follow?*

Any internal naming conventions defined in the project. No special naming conventions defined beyond those published in related papers to be released.

*Will search keywords be provided that optimize possibilities for re-use?*

N/A.

*Do you provide clear version numbers?*

N/A.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

Metadata is generated intended to be used internally within the project, as needed by other modules to reuse the results of the modules in charge of IMDEA.

### 2.1.5. Universidad Carlos III de Madrid

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

No. The data we plan to collect and process is aggregated data.

*What naming conventions do you follow?*

To be defined.

*Will search keywords be provided that optimize possibilities for re-use?*

The processed data is offered through an API to third parties which help them to make the search for the value of the specific audience they are interested in.

*Do you provide clear version numbers?*

Yes, we identify the different versions of the software with a version number.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

During the processing phase we can create metadata that allows to identify for instance ranges of value (min, max, median value), dynamics of the value evolution, etc.

### 2.1.6. Telefónica Investigación y Desarrollo

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

No.

*What naming conventions do you follow?*

To be defined.

*Will search keywords be provided that optimize possibilities for re-use?*

To be defined.

*Do you provide clear version numbers?*

No.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

Since the aim is just to create a data portability tool for the user's benefit only, there is no metadata created.

### 2.1.7. Fastweb

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

To be defined.

*What naming conventions do you follow?*

To be defined.

*Will search keywords be provided that optimize possibilities for re-use?*

To be defined.

*Do you provide clear version numbers?*

To be defined.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

To be defined.

### 2.1.8. LSTech ESPANA

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

LSTECH provides mechanisms for analysing the personal data that is collected/ stored for the project. Metadata are also available for these data. All these stored in the project platform “space”, not available for external entities, not discoverable by any means.

*What naming conventions do you follow?*

No special naming conventions. LSTECH uses the data as they are defined/ stored in the project. The naming is related to the context in order to facilitate better understanding and reusability of the data.

*Will search keywords be provided that optimize possibilities for re-use?*

N/A.

*Do you provide clear version numbers?*

In the services, applications, data, yes, where required, but not necessarily.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

Metadata is generated intended to be used internally within the project, as needed by other modules to use the results of the modules in charge of LSTECH.
Metadata standards do exist for some of the data that is collected/ used in the project, like personal contact information, and they are used for the project. We do not know if we will need to create more in this stage.

### 2.1.9. KU Leuven – CiTiP

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

Deliverables of KU Leuven – CiTiP include keywords and abbreviations that help navigating the research data in .docx and .pdf files.

*What naming conventions do you follow?*

N/A.

*Will search keywords be provided that optimize possibilities for re-use?*

Yes.

*Do you provide clear version numbers?*

Yes.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

KU Leuven – CiTiP uses keywords and abbreviations that help navigating the research data in .docx and .pdf files.

**2.1.10.Asociación de Usuarios de Internet**

*Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?*

No, the identification data is stored in encrypted form and the rest is stored disaggregated in different spaces in the database.

*What naming conventions do you follow?*

The data is stored in a MySQL database and the only limitation to the names is the implicit use of this tool.

*Will search keywords be provided that optimize possibilities for re-use?*

No.

*Do you provide clear version numbers?*

Yes.

*What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

We do not use metadata.

**2.1.11.INTERACTIVE ADVERTISING BUREAU SPAIN**

*Not applicable.*

### 2.1.12.WIBSON

*Not applicable.*

### 2.1.13.TAPTAP

*Not applicable.*

### b. Making data openly accessible

### 2.2.1. Politecnico di Torino

*Which data produced and/or used in the project will be made openly available as the default?*

Part or the whole data produced by POLITO in the project might be made openly public for its utilization. Since the produced data is aggregated data, and would not have any implication in individuals' data privacy. No data collected through the EasyPIMS demonstrator is made publicly available.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

Part or the whole data produced by POLITO might be considered private and only shared under restrictions. All EasyPIMS data falls in this category.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

Considering web archives, we could make (part of) it available for download on POLITO internal servers.

*What methods or software tools are needed to access the data?*

API is based on web technologies.

*Is documentation about the software needed to access the data included?*

Yes, all documentation to use the API has been released by the project.

*Is it possible to include the relevant software (e.g. in open-source code)?*

Yes, the PDK includes documented examples on how to access the API.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

We use the ZENODO open repository.

*Have you explored appropriate arrangements with the identified repository?*

ZENODO has been selected by the consortium.

*If there are restrictions on use, how will access be provided?*

No.

*Is there a need for a data access committee?*

No.

*Are there well described conditions for access (i.e. a machine-readable license)?*

No.

*How will the identity of the person accessing the data be ascertained?*

To be defined once the repository of the project is set up. In principle, the usual ones based on accounts/userids. Part of the data may be open accessible.

### 2.2.2. NEC Laboratories Europe

*Which data produced and/or used in the project will be made openly available as the default?*

No data.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

We have personal data, so, we are not sharing it.

*Other questions are not applicable.*

### 2.2.3. Ermes Cyber Security

*Which data produced and/or used in the project will be made openly available as the default?*

Data summarizing websites collected using web scrapers is not made available by ECS, while part or the Privacy Metrics produced by ECS in the project might be made public for its utilization using specific APIs. Sharing these data does not have any privacy implication.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

Data collected by ECS using web scrapers represent a competitive advantage for ECS, and as such these data are not shared or made publicly available. Privacy Metrics generated by ECS might be made available under restrictions as ECS plans to exploit them commercially.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

Privacy Metrics is made accessible through the use of specific APIs.

*What methods or software tools are needed to access the data?*

Whatever software tool capable of interacting with web APIs is fine.

*Is documentation about the software needed to access the data included?*

Yes, all documentation to use the APIs is released by the project.

*Is it possible to include the relevant software (e.g. in open-source code)?*

Yes, the PDK includes documented examples on how to access the APIs.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

This has not been defined yet.

*Have you explored appropriate arrangements with the identified repository?*

This has not been defined yet.

*If there are restrictions on use, how will access be provided?*

Privacy Metrics made available by ECS in the context of the project will have not access restrictions.

*Is there a need for a data access committee?*

No.

*Are there well described conditions for access (i.e. a machine-readable license)?*

No.

*How will the identity of the person accessing the data be ascertained?*

We use accounts defined by the repository hosting platform.

### 2.2.4. IMDEA Networks

*Which data produced and/or used in the project will be made openly available as the default?*

As data from IMDEA Networks' modules is intended to be reused internally in the project and/or solution, it is not openly available. The software is part of the PDK, and thus is made public as part of the deliverables of the project.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

N/A.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

No data is made accessible. In case any demo is required we link to the open public sources were data for the demo or example is downloadable.

*What methods or software tools are needed to access the data?*

N/A.

*Is documentation about the software needed to access the data included?*

There is, consequently, no specific tool to access any data, as no data is shared. However, as part of PIMCITY deliverables, we make available open-source code and its related documentation for third parties to reuse. Part of this Open-Source code are libraries from IMDEA.

*Is it possible to include the relevant software (e.g. in open-source code)?*

As part of PIMCITY deliverables, we make available open-source code and its related documentation for third parties to reuse. Part of this Open-Source code are libraries from IMDEA.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

The project website and potentially some websites like Github (to be explored) for open-source shared code.

*Have you explored appropriate arrangements with the identified repository?*

We assume that data from the project comes from the project repository. It is internal from the project, hosted by Fastweb.

*If there are restrictions on use, how will access be provided?*

Not defined.

*Is there a need for a data access committee?*

Not defined.

*Are there well described conditions for access (i.e. a machine-readable license)?*

Not defined.

*How will the identity of the person accessing the data be ascertained?*

Not defined.

### 2.2.5. Universidad Carlos III de Madrid

*Which data produced and/or used in the project will be made openly available as the default?*

This is still to be decided. Part or the whole data produced by UC3M in the project might be made openly public for its utilization.
Since the produced rata is based on aggregated data, this does not have any implication in individuals' data privacy.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

This is still to be decided. Part or the whole data produced by UC3M might be considered private and only shared under restrictions. This depends on the commercial value of the produced data.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

In principle data is accessible through an API, although internally in the consortium alternative sharing methods can be discussed.

*What methods or software tools are needed to access the data?*

As indicated above, the data is accessible through an API. The API can be queried.

*Is documentation about the software needed to access the data included?*

Yes, documentation is made available once the software is developed.

*Is it possible to include the relevant software (e.g. in open-source code)?*

The software is integrated within the PIMCITY infrastructure wherever it is determined in the design phase.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

In principle, we make it available in the repository specified by the project.

*Have you explored appropriate arrangements with the identified repository?*

We are part of the consortium, so the arrangement is granted by the existence of the project.

*If there are restrictions on use, how will access be provided?*

No restrictions.
*Is there a need for a data access committee?*

No.

*Are there well described conditions for access (i.e. a machine-readable license)?*

Once the repository from the project is set up, such conditions should be defined by third parties' access.

*How will the identity of the person accessing the data be ascertained?*

To be defined once the repository of the project is set up. In principle, the usual ones based on accounts/userids.

### 2.2.6. Telefónica Investigación y Desarrollo

*Which data produced and/or used in the project will be made openly available as the default?*

No data is made openly available by default.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

Our tasks do not involve data sharing.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

No data is made accessible since there is no data sharing.

*What methods or software tools are needed to access the data?*

Data is accessed by authorized parties only using the appropriate API that the 3rd party PIM provides. In cases that there is no API available, manual process might be considered on a case-by-case scenario.

*Is documentation about the software needed to access the data included?*

No data access is provided, and thus no documentation is required.

*Is it possible to include the relevant software (e.g. in open-source code)?*

As part of PIMCITY deliverables, we make available open-source code and its related documentation for third parties to reuse if authorized.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

As part of PIMCITY deliverables, we make available open-source code and its related documentation for third parties to reuse.

*Have you explored appropriate arrangements with the identified repository?*

We are part of the consortium, so the arrangement is granted by the existence of the project.

*If there are restrictions on use, how will access be provided?*

No restrictions.

*Is there a need for a data access committee?*

No.

*Are there well described conditions for access (i.e. a machine-readable license)?*

To be defined by the Consortium.

*How will the identity of the person accessing the data be ascertained?*

No data is committed to the repository.

### 2.2.7. Fastweb

*Which data produced and/or used in the project will be made openly available as the default?*

To be defined by the Consortium.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

To be defined. Personal traffic data from Fastweb customers cannot be shared, in compliance with privacy regulations, and we do not plan to use it in the project.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

To be defined by the Consortium.

*What methods or software tools are needed to access the data?*

To be defined by the Consortium parties developing PIMCity modules.

*Is documentation about the software needed to access the data included?*

To be defined by the Consortium parties developing PIMCity modules.

*Is it possible to include the relevant software (e.g. in open-source code)?*

To be defined by the Consortium parties developing PIMCity modules.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

The data hosted in Fastweb's cloud computing infrastructure is deposited in Fastweb data centres located in Italy.

*Have you explored appropriate arrangements with the identified repository?*

N/A.

*If there are restrictions on use, how will access be provided?*

To be defined by the Consortium parties developing PIMCity modules.

*Is there a need for a data access committee?*

To be clarified.

*Are there well described conditions for access (i.e. a machine-readable license)?*

To be defined by the Consortium parties developing PIMCity modules.

*How will the identity of the person accessing the data be ascertained?*

To be defined by the Consortium parties developing PIMCity modules.

### 2.2.8. LSTech ESPANA

*Which data produced and/or used in the project will be made openly available as the default?*

N/A.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

N/A.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

No data is made accessible.

*What methods or software tools are needed to access the data?*

N/A.

*Is documentation about the software needed to access the data included?*

N/A.

*Is it possible to include the relevant software (e.g. in open-source code)?*

N/A.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

The project website and potentially some websites like Github (to be explored) for open-source shared code.

*Have you explored appropriate arrangements with the identified repository?*

We assume that data from the project comes from the project repository. It is internal from the project, hosted by Fastweb.

*If there are restrictions on use, how will access be provided?*

Not defined.

*Is there a need for a data access committee?*

Not defined.

*Are there well described conditions for access (i.e. a machine-readable license)?*

Not defined.

*How will the identity of the person accessing the data be ascertained?*

Not defined.

### 2.2.9. KU Leuven – CiTiP

*Which data produced and/or used in the project will be made openly available as the default?*

Research data, i.e. KU Leuven – CiTiP deliverables such as legal requirements.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

KU Leuven – CiTiP does not work with datasets.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

Primarily via project website.

*What methods or software tools are needed to access the data?*

N/A.

*Is documentation about the software needed to access the data included?*

No.

*Is it possible to include the relevant software (e.g. in open-source code)?*

N/A.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

N/A.

*Have you explored appropriate arrangements with the identified repository?*

N/A.

*If there are restrictions on use, how will access be provided?*

N/A.

*Is there a need for a data access committee?*

No.

*Are there well described conditions for access (i.e. a machine-readable license)?*

N/A.

*How will the identity of the person accessing the data be ascertained?*

N/A.

### 2.2.10. Asociación de Usuarios de Internet

*Which data produced and/or used in the project will be made openly available as the default?*

The data of the users participating in the project are not made public.

*If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

Contact details of users participating in the project are not shared. They are only used by UAI to contact the user if required at any stage of the project and to keep him/her informed of the project's progress.

*Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.*

*How will the data be made accessible (e.g. by deposition in a repository)?*

No.

*What methods or software tools are needed to access the data?*

Access to the database and knowledge of encryption keys.

*Is documentation about the software needed to access the data included?*

No.

*Is it possible to include the relevant software (e.g. in open-source code)?*

No.

*Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

The data is stored in the AUI's servers which are located in a datacenter operated by the hosting provider ARSYS.

*Have you explored appropriate arrangements with the identified repository?*

No.

*If there are restrictions on use, how will access be provided?*

No.

*Is there a need for a data access committee?*

No.

*Are there well described conditions for access (i.e. a machine-readable license)?*

Yes.

*How will the identity of the person accessing the data be ascertained?*

Only people with system administrator status can access the databases and if they do, the identification data is always stored in encrypted form.

#### 2.2.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*Not applicable.*

#### 2.2.12. WIBSON

*Not applicable.*

#### 2.2.13. TAPTAP

*Not applicable.*

### c. Making data interoperable

#### 2.3.1. Politecnico di Torino

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

The data interoperability is granted by the definition and usage of open API.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

The consortium defined custom vocabularies and methodologies to make data interoperable.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

No.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

No.

### 2.3.2. NEC Laboratories Europe

*Not applicable.*

### 2.3.3. Ermes Cyber Security

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

This needs to be defined.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

This needs to be defined.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

This needs to be defined.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

Yes, if needed.

### 2.3.4. IMDEA Networks

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

Yes. We will be using Python and open software and libraries, as well as data formats to produce the results of the investigation. Even we plan to share such data and developments with papers produced throughout the project.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

Not defined.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

Not defined.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

To be clarified.

### 2.3.5. Universidad Carlos III de Madrid

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

We are not aware of other datasets of the nature of the one to be produced by UC3M. So, interoperability will need to be explored once datasets of similar nature are identified. If this happens during the execution of the project, we will analyse this aspect.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

N/A.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

N/A.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

There is a clear documentation associated to the developed tool that allows anyone with the appropriate permission to access the produced data.

### 2.3.6. Telefónica Investigación y Desarrollo

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

There is no data sharing available.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

There is no data sharing available.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

There will be no data sharing available.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

N/A.

### 2.3.7. Fastweb

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

To be defined by the Consortium parties developing PIMCity modules.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

To be defined by the Consortium parties developing PIMCity modules.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

To be defined by the Consortium parties developing PIMCity modules.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

To be clarified.

### 2.3.8. LSTech ESPANA

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

It is a decision and action of the whole consortium. At the moment LSTECH does not have specific plans to produce data to be exchanged.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

Not defined.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

Not defined.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

N/A.

### 2.3.9. KU Leuven – CiTiP

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

Yes. KU Leuven – CiTiP deliverables which include research data are easily accessible with word processing programmes.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

N/A.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

N/A.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

N/A.

### 2.3.10. Asociación de Usuarios de Internet

*Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

To be clarified.

*What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

To be clarified.

*Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

To be clarified.

*In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

Not applicable.

### 2.3.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*Not applicable.*

### 2.3.12. WIBSON

*Not applicable.*

### 2.3.13. TAPTAP

*Not applicable.*

## d. Increase data re-use (through clarifying licences)

### 2.4.1. Politecnico di Torino

*How will the data be licensed to permit the widest re-use possible?*

The data is offered for free with an unlimited usage style license.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

Data is either already available, or will never be made available to not violate the privacy constraints of end users.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

No.

*How long is it intended that the data remains re-usable?*

For at least 5 years.

*Are data quality assurance processes described?*

See 2.4.9. KU Leuven – CiTiP

### 2.4.2. NEC Laboratories Europe

*Not applicable.*

### 2.4.3. Ermes Cyber Security

*How will the data be licensed to permit the widest re-use possible?*

This will be defined later in the project.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

It is too early to define deadlines at this state of the project.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

Since ECS is a commercial company, Privacy Metrics generated by ECS within the project might be usable by third parties under restrictions (technical, commercial and legal) which will be defined later in the project.

*How long is it intended that the data remains re-usable?*

To be defined.

*Are data quality assurance processes described?*

To be defined.

### 2.4.4. IMDEA Networks

*How will the data be licensed to permit the widest re-use possible?*

No data will be provided, except the data and algorithms useful to reproduce the research conducted by IMDEA.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

N/A.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

N/A.

*How long is it intended that the data remains re-usable?*

N/A.

*Are data quality assurance processes described?*

N/A.

### 2.4.5. Universidad Carlos III de Madrid

*How will the data be licensed to permit the widest re-use possible?*

This will be defined later in the project once we have more information and a better judgment to make the decision. For instance, if the data has commercial value, it will be probably licensed through commercial licenses. Another option is to license the data in a free manner. Finally, a combination of the two options is also foreseen.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

The data is only made available for re-use once the internal goals of UC3M and PIMCITY that requires the use of data in exclusivity are achieved.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

Yes, the data is usable by third parties further than the end of the project. This use is ruled through the specific license that is decided to apply to the data.

*How long is it intended that the data remains re-usable?*

This is hard to predict. If the data produced shows a clear commercial value, it remains re-usable for an undefined period of time until it loses its commercial value. Similar principles apply for research use of the data.

*Are data quality assurance processes described?*

We conducted data quality assurance tests and properly reported when it was/is due.

### 2.4.6. Telefónica Investigación y Desarrollo

*How will the data be licensed to permit the widest re-use possible?*

No data is shared.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

No data is shared.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

No data is shared.

*How long is it intended that the data remains re-usable?*

No data is shared.

*Are data quality assurance processes described?*

No data is shared.

### 2.4.7. Fastweb

*How will the data be licensed to permit the widest re-use possible?*

To be defined by the Consortium parties developing PIMCity modules.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

To be defined by the Consortium parties developing PIMCity modules.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

To be defined by the Consortium parties developing PIMCity modules.

*How long is it intended that the data remains re-usable?*

To be defined by the Consortium parties developing PIMCity modules.

*Are data quality assurance processes described?*

To be defined by the Consortium parties developing PIMCity modules.

#### 2.4.8. LSTech ESPANA

*How will the data be licensed to permit the widest re-use possible?*

No data is provided.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

N/A.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

N/A.

*How long is it intended that the data remains re-usable?*

N/A.

*Are data quality assurance processes described?*

N/A.

#### 2.4.9. KU Leuven – CiTiP

*How will the data be licensed to permit the widest re-use possible?*

KU Leuven deliverables are be licensed.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

As soon as possible; see general approach to data management as identified in the data management plan.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

See general approach to data management as identified in the data management plan.

*How long is it intended that the data remains re-usable?*

See general approach to data management as identified in the data management plan.

*Are data quality assurance processes described?*

See general approach to data management as identified in the data management plan.

### 2.4.10. Asociación de Usuarios de Internet

*How will the data be licensed to permit the widest re-use possible?*

N/A.

*When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

N/A.

*Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

No.

*How long is it intended that the data remains re-usable?*

N/A.

*Are data quality assurance processes described?*

N/A.

### 2.4.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*Not applicable.*

### 2.4.12. WIBSON

*Not applicable.*

### 2.4.13. TAPTAP

*Not applicable.*

## 3. ALLOCATION OF RESOURCES

*What are the costs for making data FAIR in your project?*

Costs are mostly represented by two main contributions:
1. Costs to manage and make data accessible for a Fair approach and
2. Costs for hosting data API in servers that hosts data and offer API to access it.

*How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).*

Each partner is responsible for the costs incurred in making data FAIR, and for open access.

*Who will be responsible for data management in your project?*

Responsible partners are indicated in the table at the beginning of this document. In particular, specific roles are foreseen for technical coordinator (T8.2); dissemination manager (T6.1); innovation manager (T6.4); data manager (T7.1). Besides, PIMCity data protection officer team, subject to their competences, advised the PIMCity project partners on data management.

*Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?*

To be defined and agreed in the future.

## 4. DATA SECURITY

### 4.1. Politecnico di Torino

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

POLITO does not store or transfer sensitive data in the context of PIMCITY.
In general, all data collected and produced by POLITO is stored in internally hosted datacentres, protected by standard mechanisms like firewalls. The access to the servers and services is based on personal authentication via credentials. Each access is logged.

*Is the data safely stored in certified repositories for long term preservation and curation?*

No – and likely not in the near future.

### 4.2. NEC Laboratories Europe

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

All data is transferred over encrypted connections.

*Is the data safely stored in certified repositories for long term preservation and curation?*

The data is stored at NEC premises. The computer where the data is stored is located in a room where only authorized personnel can access. The computer where the data is stored does not have direct connection to the Internet.

### 4.3. Ermes Cyber Security

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

In general, ECS follows rather strict policy for data security. See DPIA for details. In general,
• all data collected and produced by ECS are stored and transferred using strong encryption;
• data are stored on both internal servers and private remote datacentres in multiple copies for redundancy;
• all accesses are protected by firewalls, logged and allowed to authorized personnel only.

*Is the data safely stored in certified repositories for long term preservation and curation?*

Yes, all data collected by ECS are safely duplicated in private remote datacentres.

### 4.4. IMDEA Networks

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

No sensitive data is stored or transferred. We only use open data available in the internet without any "personal" information.

*Is the data safely stored in certified repositories for long term preservation and curation?*

It is stored in the computers used in the research process.

### 4.5. Universidad Carlos III de Madrid

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

The data is protected using start-of-the-art security practices. It is stored in servers at UC3M which provide a full protected network with firewall, traffic monitoring, server access control (user, passwords, encryption options), servers located in rooms with physical access control, periodic data backups, etc.

If the data is moved to a different repository based on project's decision, similar security provisions are demanded.

*Is the data safely stored in certified repositories for long term preservation and curation?*

Not that we are aware of.

### 4.6. Telefónica Investigación y Desarrollo

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

All data is stored encrypted using state-of-the-art encryption technology. Even if the stored data gets lost, the user can request another data importation from the original sources and restore the data.

*Is the data safely stored in certified repositories for long term preservation and curation?*

No, we do not store data.

### 4.7. Fastweb

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

Fastweb's cloud infrastructure has all the security, privacy and continuity provisions included in the international standards for information security (ISO 27001), business continuity (ISO 22301), IT service management (ISO 20000), cloud services security (ISO 27017), cloud service privacy (ISO 27018), and security incident management (ISO 27035). Fastweb holds certifications for the aforementioned standards, issued by accredited third parties.

*Is the data safely stored in certified repositories for long term preservation and curation?*

To be clarified.

### 4.8. LSTech ESPANA

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

LSTECH is not storing data for the project at its own premises. No sensitive data is stored or transferred.

To be clarified.

*Is the data safely stored in certified repositories for long term preservation and curation?*

It is stored in the computers used in the research process.

To be clarified.

### 4.9. KU Leuven – CiTiP

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

To be clarified.

*Is the data safely stored in certified repositories for long term preservation and curation?*

To be clarified.

### 4.10. Asociación de Usuarios de Internet

*What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

N/A.

*Is the data safely stored in certified repositories for long term preservation and curation?*

N/A.

### 4.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*Not applicable.*

### 4.12. WIBSON

*Not applicable.*

### 4.13. TAPTAP

*Not applicable.*

## 5. ETHICAL ASPECTS

*Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).*

Please see deliverables D7.1, D7.2, D9.1, D9.2.

### 5.1. Politecnico di Torino

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

The EasyPIMS platform has been designed and follows the best practice to collect the explicit end user consent.

### 5.2. NEC Laboratories Europe

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

No.

### 5.3. Ermes Cyber Security

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

We did not collect personal data in the context of the project.

### 5.4. IMDEA Networks

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

N/A.

### 5.5. Universidad Carlos III de Madrid

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

We do not plan to collect any personal data.

### 5.6. Telefónica Investigación y Desarrollo

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

No.

### 5.7. Fastweb

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

As of April 2020, Fastweb does not plan to collect personal data.

### 5.8. LSTech ESPANA

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

N/A.

### 5.9. KU Leuven – CiTiP

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

N/A.

### 5.10. Asociación de Usuarios de Internet

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

Consent is requested on all forms where users are required to provide personal data and are informed of the use that will be made of it and their rights in accordance with the GDPR.

### 5.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

N/A

### 5.12. WIBSON

*Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?*

N/A.

### 5.13. TAPTAP

*Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?*

N/A.

## 6. OTHER

### 6.1. Politecnico di Torino

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

We follow the best practice that are commonly used in the research community.

### 6.2. NEC Laboratories Europe

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

We follow the IMSS policies of NEC.

### 6.3. Ermes Cyber Security

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

ECS is completing the process to get ISO 9001 and 27001 certifications.

### 6.4. IMDEA Networks

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

To be clarified.

### 6.5. Universidad Carlos III de Madrid

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

We follow best practices from the research community for data collection, processing and management.

### 6.6. Telefónica Investigación y Desarrollo

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

Any use/access/processing of data follows internal departmental policies of TID.

### 6.7. Fastweb

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

We follow industry best practices and comply with Italian laws concerning the generation, handling and deletion of data.

### 6.8. LSTech ESPANA

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

N/A.

### 6.9. KU Leuven – CiTiP

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

To be clarified.

### 6.10. Asociación de Usuarios de Internet

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

To be clarified.

### 6.11. INTERACTIVE ADVERTISING BUREAU SPAIN

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

N/A.

### 6.12. WIBSON

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

N/A.

### 6.13. TAPTAP

*Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?*

N/A.