

# D 6.3 Data Management Plan - Final

<b>Document type</b>	Report
<b>Deliverable N°</b>	D 6.3
<b>Workpackage</b>	WP6
<b>Leading partner</b>	CES
<b>Author(s)</b>	Michelangelo Secchi, Bruno Sousa
<b>Dissemination level</b>	Public
<b>Status</b>	Version for final review
<b>Date</b>	22 May 2018
<b>Version</b>	1.0

## **Executive Summary**

This Data Management Plan describes the data management life cycle for the data collected, processed and generated by the EMPATIA project.

Following the guidelines provided by the EC on the Horizon2020 Manual, this DMP aims to make research results findable, accessible, interoperable and re-usable (FAIR), and includes details regarding:

- the handling of research data during & after the end of the project
- data collected, processed and generated in each dataset
- methodology & standards applied
- means for ensuring Open Access (when applicable)
- means of preservation after the end of the project

## Table of Contents

Executive Summary.....	2
1. Introduction .....	5
1.1. Purposes of Data Management Plan.....	5
1.2. WPs and Tasks of EMPATIA Related to the DMP .....	5
2.1. Noted Model for Dataset Description.....	6
3. EMPATIA Dataset List .....	9
3.1. Dataset #1 – Social Ethical and Legal Context of Pilots .....	11
3.2. Dataset #2 – Platform Ethics SoA .....	14
3.3. Dataset #3 – Platforms’ performance data .....	18
3.4. Dataset #4 – Scientific Dissemination .....	21
3.5. Dataset #5 – Dissemination.....	24
3.6. Dataset #6 – Mapping Participatory Innovations Survey.....	27
3.7. Dataset #7 – Evaluation & Impact Assessment .....	30
3.8. Dataset #8 – Pilot - Lisbon .....	35
3.9. Dataset #9 – Pilot - Milan .....	39
3.10. Dataset #10 – Pilot – Říčany .....	43
3.11. Dataset #11 –Pilot – Wuppertal.....	46

## Acronyms

Acronyms	Description
Action	The Action entitled 'Enabling Multichannel Participation Through ICT Adaptations — EMPATIA' ('action'), as described in the DoA.
CA	Consortium Agreement
CES	Centro de Estudos Sociais – University of Coimbra (Leader of EMPATIA)
CML	Camara Municipal Lisboa (Municipal Chamber of Lisbon, Portugal)
D21	Democracy 21 (Partner of EMPATIA)
D n.n	Deliverable (referred to EMPATIA's proposal)
DDI	Digital Democratic Innovations
DI	Democratic Innovation
DMP	Data Management Plan
DoA	Description of Action, detailed description of the EMPATIA project annexed to the GA.
EC	European Commission
GA	Grant Agreement
H2020	Horizon 2020
INE	Instituto Nacional de Estatística (National Statistical Institute of Portugal)
IL	IN Loco (Partner of EMPATIA)
IP	Intellectual Property (often used also as IPR – Intellectual Property Rights)
M nn	Month (numbered since the beginning of the Action: M01 = Jan 2016)
NUTS	Nomenclature of Territorial Units for Statistics
OA	Open Access
ONE	OneSource Ltd. (Partner of EMPATIA, Technical Coordinator)
PB	Participatory Budgeting
PS	Partido Socialista – Socialist Party, Portugal
SaaS	Software as a Service
SEL	Social, ethical and Legal (analysis)
T n.n	Task (referred to EMPATIA's proposal)
UNIMI	Università degli Studi di Milano (Partner of EMPATIA)
WP n	Working Package (referred to EMPATIA's proposal)
ZLOG	ZebraLog (Partners of EMPATIA, Responsible for Pilots)

# 1. Introduction

The Data Management Plan (DMP) follows the H2020 guidelines published on version 3 of FAIR Data Management in Horizon 2020. The FAIR Data Management goals aim to make research data Findable, Accessible, Interoperable and Reusable (i.e. FAIR). The DMP is a key tool to achieve the FAIR goals and provides useful insights regarding the good management practices for research data. Explanation of datasets is performed according to the methodology followed in other projects, such as TANDEM and other analysed in the preparatory stage.

## 1.1. Purposes of Data Management Plan

- First, to organize and facilitate the management of research data during the delivery of the Action, ensuring effective integration between a decentralized structure of independent datasets managed locally by different partners of the EMPATIA consortium.
- Second, to comply with the provisions related to Data Management and Open Access in H2020, defined by EUROPEAN COMMISSION Directorate-General for Research & Innovation and specified in the Art 29.3 of the Grant Agreement of EMPATIA. EMPATIA is committed to ensure transparency and access to research data generated under the Action.
- Third to ensure the compliance of the management of data during the Action with the social, ethical and legal principles established by the EMPATIA Consortium, with a focus on the protection of personal data of users of EMPATIA's platform during the Pilots.

## 1.2. WPs and Tasks of EMPATIA Related to the DMP

Datasets of EMPATIA are generated and managed under the organizational framework of the project, consistently with the different domains covered by each Working Package:

- **WP1** collects and generates data for a variety of research purposes, including quantitative and qualitative information collected both online and in person.
- **WP2** collects and generates technical data for validation as well as for research purposes, analyzing non-personally-identifying information.
- **WP3** manages data necessary for the implementation of pilots in the cities of Lisbon (PT) Wuppertal (GER) and Ričany (CZR) where the EMPATIA platform will be used to manage PB processes and other democratic innovations.
- **WP4** manages data for evaluation purposes applying a multilevel assessment based on the collection and analysis of various key performance indicators (Technical, Behavioural, Socio-economic, Political KPIS: inclusiveness and political alienation, and other Process related KPIS)
- **WP5** manages data for dissemination and exploitation purposes applying a multilevel assessment based on the collection and analysis of various key performance indicators (Technical, Behavioural, Socio-economic, Political KPIS: inclusiveness and political alienation, and other Process related KPIS)

## 2. Methodology

The DMP of EMPATIA integrates the different datasets generated by all WPs throughout the Action into a unique management model.

The complete list of datasets and metadata have been managed collecting data provided by all partners responsible for each of the datasets generated in EMPATIA (see Chapter 3), according to the guidelines provided.

Since the beginning of the action a Model for dataset description has been made available on the internal wiki of the consortium at [wiki.empatia-project.eu](http://wiki.empatia-project.eu)

During the Action each partner responsible for a dataset have been continuously reviewing and updating the information required according to the model provided and to the ethical guidelines defined by WP1.

The Ethical Coordinator supervised the operation of data management providing guidance in all issues related to the management and processing of personal data during the Action.

### 2.1. Noted Model for Dataset Description

The methodology followed to describe each dataset is based on the FAIR guidelines and on fields identified on data management plans of other projects, such as TANDEM.

The FAIR guidelines suggest a set of questions that can be classified into distinct groups:

- **Dataset identification**, which includes all the fields to identify uniquely the dataset, and also answer to questions like the purpose of the data collection/generation and its relation with the project's goals.
- **Dataset responsibilities**, which is a group that clearly identifies the data management responsibility for each dataset. We identify three type of entities, which have different interactions with the data:
  - **Data Controller**: Entity responsible for data management.
  - **Data Processor**: Entity responsible to process, generate and analyse the data for the datasets (according to the purposes defined for each instance)
  - **Responsible for Data Storage**: Entity responsible for storing and securing the data in the dataset.

Each dataset created under EMPATIA will include a clear definition of how these three level of responsibility are attributed to partners of the consortium and/or to third parties depending on the nature and purpose of the data collection.

- **Data Collected and Standards**, group that identifies the data collected and the standards employed in to their management. It should be noticed that this group is relevant to make the data findable. We refer to the following categories of data:
  - **Personal Data**: It refers to the personal data required to participate to the services and research activity promoted by EMPATIA online and in person. (e.g. name, email address, age, gender, etc.). Personal data collected for authentication purposes are generally mandatory in pilots, while the majority of the others personal data are collected on voluntary basis.

- **Votes and Preferences:** Votes and preferences are core functions of collaborative platforms for civic engagement, generated by the interaction between users and ongoing discussions.
  - **User Generated Content:** Any public content (text or any other media) generated by users through their activity on the platform. In the Pilots of EMPATIA, it mainly refers to ideas/proposals/projects (text and any other media) generated and developed within each consultation process managed through EMPATIA..
  - **Surveys and other Personal data voluntarily generated by users:** It refers to the data generated by the users who decides to answer to questionnaires and surveys upon request of EMPATIA, online and in person. (e.g. surveys on users satisfaction, questionnaire on trust in politics, etc.)
  - **Non-personally identifiable information:** It refers to non-personally-identifying information of the sort that web browsers and servers typically make available, such as the browser type, language preference, referring site, and the date and time of each visitor request, collected through the use of the IT services of EMPATIA.
  - **Potentially personally-identifiable information:** It refers to potentially personally-identifying information like Internet Protocol (IP) addresses, collected through the use of the IT services of EMPATIA.
- **Data Sharing**, which is a group that points out the methods, standards employed to share the data of the dataset.
  - **Archiving and preservation** group that presents the archiving policies, periods and mechanisms of preserving the data.
  - **Data Security** group that delineates the logical and physical mechanisms of security that are employed to secure the access of data in the dataset.
  - **Ethical remarks** group that identifies the ethical risks associated with the dataset and the legal/regulation mechanisms that should be applied to protect data of users.

The following Table provides a template to describe the structure of a dataset according to the guidelines of the FAIR Data Management. In the following chapter each Dataset generated under EMPATIA will be described and analyzed according to the metadata described in the following Table.

Dataset Identification	
Dataset Description	Short explanation of the contents that can be found in the dataset
Dataset Purpose	Purpose of Data Collection
Dataset Source	Where the dataset is available

Dataset Date/Version	Date on which the dataset has been compiled or documented and information regarding the version of the dataset (if available)
Dataset Language	The default language of the data that can be found in the dataset
Related WP(s) and task(s)	WP(s) and respective task(s) that produce the dataset.
<b>Data responsibilities</b>	
Data Processor	Entity/Person responsible to process the data in the dataset
Data Controller	Entity/Person responsible to manage the dataset
Responsible for Data Storage	Entity/Person responsible for securing the data in the dataset
<b>Data Collected and Standards</b>	
List of Data Collected	Vocabularies of metadata that are required to gather data or to make it interoperable
Standards (formats, estimated volume of data)	Employed standards, formats to collect, store data
<b>Data Sharing</b>	
Data Shared	Labels of fields to be shared (see former section)
Data sharing purpose (for dissemination or exploitation)	Reasons to proceed with data sharing, either for dissemination or exploitation.
Format of data sharing	What is the format of shared data
Means of data sharing	Which means are employed to share data
Embargo Period	Eventual embargo period
Open Data license	If data is shared with open access and what is the license
<b>Archiving and preservation</b>	
Data Storage (including backups)	Storage policies associated with the dataset, including backups and their frequency
Preservation periods and where	Preservation periods associated with the backups (e.g. if available after the conclusion of the project) and the location where they are stored (i.e. in a datacentre).
<b>Data Security</b>	
Physical protection mechanisms	Physical protection mechanisms to secure data, including restricted access to datacentres, authentication mechanisms.
Logical protection mechanisms	Logical protection mechanisms to secure data, including encryption, authentication and privacy mechanisms.
<b>Ethical remarks</b>	
Personal data protection	If personal data is included in the dataset and description of the security mechanisms to assure that such data is not accessible, not disclosed and which policies apply to removed users.
Legal framework	If dataset has associated any legal requirements (National, European laws)
Policies /Information Sheet	If there are EMPATIA's policies applicable to this dataset
Ethical risks	Ethical risks that may be present in the dataset

### 3. EMPATIA Dataset List

The Datasets references have been defined by EMPATIA partners considering the EMPATIA use cases, the pilots and surveys regarding platform usage, users and or organizers, such list is provided in the following Table. The list of datasets includes the responsible partner and identifies the WP where the data is generated. It should also be noticed that the provided list is not static and may be modified (addition/removal of datasets) according to the project developments. The next version of the EMPATIA DMP will describe such modifications if they occur, especially for other data sources for instance new or not planned pilots.

This section identifies the datasets generated in EMPATIA and the entities that generate or use the data present in such datasets.

Name	Description	WP	Repository Link	Ref
Dataset #1 – Social Ethical and Legal Context of <b>Pilots</b>	This dataset includes public a data regarding the context of implementation of the four pilots of EMPATIA, collected thorough secondary sources. This dataset is used to deliver the of Social, Ethical, and Legal analysis of EMPATIA's pilots	1	<a href="https://dkan.empatia-project.online/dataset/31-dataset-1-%E2%80%93-social-ethical-and-legal-context-pilots">https://dkan.empatia-project.online/dataset/31-dataset-1-%E2%80%93-social-ethical-and-legal-context-pilots</a>	CES
Dataset #2 – Platform Ethics	This dataset is the result of the Ethical analysis of a sample of sixteen collaborative platforms currently used in the management of DIs in Europe, Latin America, Canada and United States. The analysis focused on the state of the art and aims to collect input and define ethical standards to be adopted in the deployment of EMPATIA.	1	<a href="https://dkan.empatia-project.online/dataset/empatia-db-02-platforms-soa">https://dkan.empatia-project.online/dataset/empatia-db-02-platforms-soa</a>	CES
Dataset #3 – Platforms' performance data	This dataset includes the information regarding the performance of the EMPATIA platform.	2	NA	ONE
Dataset #4 – Scientific Dissemination	This dataset includes the information regarding the Scientific Dissemination of the project, including Scientific Articles submitted and approved for publications, books and chapters in books, papers presented to scientific conferences, articles in specialized magazines, and any other scientific result disseminated through the scientific community.	5	<a href="https://dkan.empatia-project.online/dataset/dataset-4-%E2%80%93-scientific-dissemination">https://dkan.empatia-project.online/dataset/dataset-4-%E2%80%93-scientific-dissemination</a>	CES
Dataset #5 – Dissemination	This dataset includes the information and personal data collected through the main website of the project <a href="http://empatia-project.eu/">http://empatia-project.eu/</a> . The main website will be used to disseminate the result of EMPATIA to the broader public possible and at the same time to feed and maintain a community of users and followers of EMPATIA.	5	NA	CES
Dataset #6 – Mapping Participatory Innovations Survey	This dataset includes the information and personal data collected through the survey website of the project <a href="https://oidp.empatia-project.eu/">https://oidp.empatia-project.eu/</a> . The survey website is sponsored by the OIDP network, the OIDP 2017 conference, the EMPATIA project, and Participedia project. The objective of this survey is to collect and share data about a variety of participatory processes implemented by OIDP partners, and at the same time to feed	1	<a href="https://dkan.empatia-project.online/dataset/dataset-6-%E2%80%93-mapping-participatory-innovations-survey">https://dkan.empatia-project.online/dataset/dataset-6-%E2%80%93-mapping-participatory-innovations-survey</a>	CES

	and maintain a community of users and followers of democratic innovations which are related to EMPATIA.			
Dataset #7 – Evaluation & Impact Assessment	This dataset includes the data collected for the purposes of the evaluation and impact assessment of the project. It includes data collected via surveys delivered both online and in-person to stakeholders and participants of EMPATIA's pilots.	4	<a href="https://dkan.empatia-project.online/dataset/dataset-7-%E2%80%93-evaluation-impact-assessment">https://dkan.empatia-project.online/dataset/dataset-7-%E2%80%93-evaluation-impact-assessment</a>	BRUNEL
Dataset #8 – Pilot - Lisbon	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Lisbon (Portugal) through the site: <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a>	3	<a href="https://dkan.empatia-project.online/dataset/empatia-db-8-pilot-lisbon">https://dkan.empatia-project.online/dataset/empatia-db-8-pilot-lisbon</a>	ONE
Dataset #9 – Pilot - Milan	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Milan (Italy) through the dedicated website <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a> . The EMPATIA's service is deployed in this pilot in integration with the services managed by the partner UNIMI, based on the open and free tool OpenDCN.	3	<a href="https://dkan.empatia-project.online/dataset/empatia-db-9-pilot-milan">https://dkan.empatia-project.online/dataset/empatia-db-9-pilot-milan</a>	Municipality of Milan
Dataset #10 – Pilot – Říčany	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Říčany (CZ Republic) through the dedicated website: <a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a> . The EMPATIA's service is deployed in this pilot in integration with the services of the partner D21, based on proprietary technology.	3	<a href="https://dkan.empatia-project.online/dataset/empatia-db-10-pilot-ri%C3%A7any">https://dkan.empatia-project.online/dataset/empatia-db-10-pilot-ri%C3%A7any</a>	D21
Dataset #11 –Pilot – Wuppertal	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Wuppertal (Germany) through the dedicated website: <a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a> .	3	<a href="https://dkan.empatia-project.online/dataset/empatia-db-11-pilot-wuppertal">https://dkan.empatia-project.online/dataset/empatia-db-11-pilot-wuppertal</a>	ZLOG

### 3.1. Dataset #1 – Social Ethical and Legal Context of Pilots

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/31-dataset-1-%E2%80%93-social-ethical-and-legal-context-pilots">https://dkan.empatia-project.online/dataset/31-dataset-1-%E2%80%93-social-ethical-and-legal-context-pilots</a>
Dataset Description	This dataset includes the information regarding the context of implementation of the four pilots of EMPATIA. This dataset is used to deliver the of Social, Ethical, and Legal analysis of EMPATIA's pilots
Dataset Purpose	To collect information regarding the Social, Ethical and Legal dimension in the context of implementation of pilots
Dataset Source	Data collected through secondary sources (indicated in the dataset).
Dataset Date/Version	31/12/2016 version 1.0
Dataset Language	English
Related WP(s) and task(s)	WP1 - Foundations and Models for Multichannel Participatory Budgeting, T1.3 Social, Ethical and Legal Analysis
Data responsibilities	
Data Processor	Centro de Estudos Sociais
Data Controller	Centro de Estudos Sociais
Responsible for Data Storage	Centro de Estudos Sociais
Data Collected and Standards	
List of Data Collected	<p><b>SOCIAL</b></p> <ul style="list-style-type: none"> <li>- Population (Resident) How many inhabitants live in the municipality?</li> <li>- Area What is the Municipal Area?</li> <li>- Age Age composition</li> <li>- GDP/pc What is the Average GDP per Capita?</li> <li>- Education Education levels</li> <li>- Employment %</li> <li>- Entity responsible What entity is responsible for the budget's approval and implementation of the outcomes of the PB process (region/province/municipality/submunicipal, other)?</li> <li>- Political Party Describe the coalition</li> <li>- Voting Turnout (local elections) What was the voting turnout at the last election?</li> <li>- Other DI Do the Municipality carry out other Dis? Check Note</li> </ul> <p><b>ETHICAL</b></p> <ul style="list-style-type: none"> <li>- Personal Data collected Personal Data collected</li> <li>- Personal Data Protection Means of Personal Data Protection</li> </ul>

	<ul style="list-style-type: none"> <li>- Information Sheet access How are Information on Data Management and Personal Data protection made accessible?</li> <li>- Cookie Is the EU cookie directive respected</li> <li>- Oblio Is there any provided procedures to comply with the right to be forgotten?</li> <li>- Trackers (Ghostery) Are there any 3d party trackers?</li> <li>- Open Data What Kind of Data regarding the process are accessible?</li> <li>- Purpose What is the main purpose of data collection?</li> <li>- Archiving Where Data are archived?</li> <li>- Preservation How long data are preserved?</li> <li>- Responsibility Who is responsible for Data processing?</li> <li>- Ownership Who is owner of Data?</li> <li>- Third Parties Are there third parties with access rights?</li> <li>- Data Access Format What is the format of release:</li> <li>- Data Access Procedure What is the procedure to access Data?</li> </ul> <p>LEGAL</p> <p>Budgetary Autonomy Is the Entity financially autonomous? (if not what other subject shall authorize budget approval and execution?)</p> <ul style="list-style-type: none"> <li>- Value of the Budget (last or avg last 3 ) What is the overall value of the budget?</li> <li>- Current/investments What is the value of the current expenditures/investments</li> <li>- Budget responsibility What institutional body is in charge of Budget approval?</li> <li>- Administrative responsibility Budget What institutional body is responsible for Budget elaboration?</li> <li>- Main competences What are the main competences of the entity involved (list)?</li> <li>- Administrative responsibility What Administrative Body (office/Sector)is in charge of process delivery?</li> <li>- Cost of PB management Overall costs (estimation)</li> <li>- Third Party Implementation Are there third party involved in the implementation of the process</li> <li>- 3rd party selection What is the mechanisms of 3rd Party selection? (tender, direct, other-explain)</li> <li>- ICT Services Who is in charge of the management of digital services connected to PB? (office/third party/other - explain)</li> <li>- ICT Services CMS Who is in charge of content management of digital services connected to PB? (office/third party/other - explain)</li> <li>- Regulatory Framework ext Is there any official act that makes outcomes of PB binding for Public Policy makers?</li> <li>- Procedural Framework int Is there any official act that regulates the procedure of PB?</li> <li>- DP Regulation What regulations enforce locally Data Protection?</li> <li>- DP Compliance Are the data protection strategies compliant with local, national,international regulations?</li> <li>- DP Authority What local/national/international authority oversee over controversies regarding the enforcement of Data Protection and other relevant Ethical Issues</li> </ul>
Standards (formats, estimated volume of data)	Formats/Standards: CSV, ODT, DOC, images (GIF). Estimated Volume: Below 200 Mbytes.

Data Sharing

Data Shared	All data collected
Data sharing purpose (for dissemination or exploitation)	For transparency purposes and to improve the quality of the Social, Ethical and Legal analysis of the pilot of EMPATIA.
Format of data sharing	Data will be released: - as charts and other info-graphic elaborations (PDF)
Means of data sharing	Published in D1.3 and D1.5 of EMPATIA
Embargo Period	NA
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved at CES, Coimbra, in independent electronic sheets.
Preservation periods and where	Data is kept for the duration of pilots and 36 months after the conclusion date of EMPATIA.
<b>Data Security</b>	
Physical protection mechanisms	NA
Logical protection mechanisms	NA
<b>Ethical remarks</b>	
Legal framework	NA
Policies /Information Sheet	NA
Ethical risks	NA

### 3.2. Dataset #2 – Platform Ethics SoA

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/empatia-db-02-platforms-soa">https://dkan.empatia-project.online/dataset/empatia-db-02-platforms-soa</a>
Dataset Description	This dataset is the result of the Ethical analysis of a sample of sixteen collaborative platforms currently used in the management of DIs in Europe, Latin America, Canada and United States. The analysis focused on the state of the art and aims to collect input and define ethical standards to be adopted in the deployment of EMPATIA.
Dataset Purpose	- To analyse the state of the art of collaborative platform according to the ethical indicator established by EMPATIA - To collect input and define ethical standards to be adopted in the deployment of EMPATIA.
Dataset Source	Data collected through: <ul style="list-style-type: none"> <li>- Secondary analysis on existing sources (Articles, books, grey literature)</li> <li>- Testing sessions on the digital tools analysed</li> </ul>
Dataset Date/Version	1.0 December 2016
Dataset Language	English
Related WP(s) and task(s)	WP1 - Foundations and Models for Multichannel Participatory Budgeting;T1.1 - Theories, models and cases studies for Participatory Budgeting.
Data responsibilities	
Data Processor	Centro de Estudos Sociais
Data Controller	Centro de Estudos Sociais
Responsible for Data Storage	Centro de Estudos Sociais
Data Collected and Standards	
List of Data Collected	<ol style="list-style-type: none"> <li>1. ID <ol style="list-style-type: none"> <li>1.1. ID</li> <li>1.2. Platform name</li> <li>1.3. Case observed</li> <li>1.4. Grant type, if any</li> </ol> </li> <li>2. Methodological/functional <ol style="list-style-type: none"> <li>2.1. What kind of deployment is it possible?</li> <li>2.2. What level of tech skill are required for every kind of deployment? (U = user, M = manager, A = admin)</li> <li>2.3. What are the software languages / technologies involved?</li> <li>2.4. What kind of OS is supported / used? (server side, the user side interaction works over web)</li> <li>2.5. What kind of web server is supported / used?</li> </ol> </li> </ol>

- 2.6. What kind of functions are delivered through the platform, with respect of EMPATIA components?
- 2.7. Core components (referred to the list of EMPATIA's components as described in D2.1 – Platform Architecture)
- 3. Ethical approach: standards
  - 3.1. What is the license?
    - 3.1.1. If open source, where the code is published?
    - 3.1.2. Where and how the software license is published?
    - 3.1.3. License accessibility evaluation, 1 to 5 (1 is non-accessible, 5 is fully accessible)
  - 3.2. Is it present any 3d party software provided with a different licence within the core platform features?
    - 3.2.1. If any, what is its purpose?
    - 3.2.2. If any, what is the licence?
    - 3.2.3. If open source, where the code is published?
    - 3.2.4. If closed source, additional remarks?
    - 3.2.5. Other, specify
  - 3.3. Is there any difference based on the chosen deployment?
- Privacy & Personal Data Protection
  - 3.4. There are Information Sheets about the platform collected and mined data?
  - 3.5. Is the EU cookie directive respected (clearly opt-out possibility from cookie 3d party tracking at first user's visit)?
  - 3.6. There is explicit informed consent request?
    - 3.6.1. Where is it published?
    - 3.6.2. Level of understandability, 1-5 (1 = only tech & legal experts; 5 = every literate folks)
  - 3.7. List of Data collected
  - 3.8. Metadata e standard
  - 3.9. Archiving and Preservation
  - 3.10. Third Parties procedures
  - 3.11. Is there any provided procedures to comply with the right to be forgotten?
  - 3.12. What kind of data are automatically collected through 3d party trackers?
  - 3.13. Security features in order to ensure Privacy & personal data protection (security standard compliance)
  - 3.14. Double-step authentication YES/NO
  - 3.15. HTTPS support YES/NO
  - 3.16. Password recovery YES/NO
  - 3.17. OAuth standard YES/NO
- Open access
  - 3.18. Accessibility standard (1-5)
  - 3.19. Data are exportable?
  - 3.20. Social media interaction?
  - 3.21. API availability?
  - 3.22. Ethical principles?
  - 3.23. Additional remarks
- 4. Legal
  - 4.1. What level of legal autonomy has the platform?
  - 4.2. Are there any difference depending on the deployment type?

	See also D1.3 - Social, ethics and legal analysis – preliminary
Standards (formats, estimated volume of data)	<ul style="list-style-type: none"> <li>- Formats/Standards: CSV, ODT, DOC, images (GIF).</li> <li>- Estimated Volume Below 500 Mbytes.</li> </ul>
<b>Data Sharing</b>	
Data Shared	All data collected
Data sharing purpose (for dissemination or exploitation)	For transparency purposes and to improve the quality of the Social, Ethical and Legal analysis of the pilot of EMPATIA.
Format of data sharing	Data will be released: <ul style="list-style-type: none"> <li>- in machine readable format (CSV)</li> <li>- as charts and other info-graphic elaborations (PDF)</li> <li>- images (GIF)</li> </ul>
Means of data sharing	Published in D1.5 - Social, ethics and legal analysis – final, and on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>
Embargo Period	NA
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved at CES, Coimbra, in independent electronic sheets.
Preservation periods and where	Data is kept for the duration of pilots and 36 months after the conclusion date of EMPATIA.
<b>Data Security</b>	
Physical protection mechanisms	NA
Logical protection mechanisms	NA
<b>Ethical remarks</b>	
Legal framework	NA
Policies /Information Sheet	NA
Ethical risks	NA



### 3.3. Dataset #3 – Platforms’ performance data

Dataset Identification	
Repository	NA
Dataset Description	This dataset includes the information regarding the performance of the EMPATIA platform.
Dataset Purpose	1- To collect and store data related with the performance of the EMPATIA platform; 2- To assess the resource consumption of the diverse components of the platform;
Dataset Source	Data collected through the running EMPATIA platforms after the respective deployments: 1- On every operation performed by each component 2- During the usage of the platform by users, administrator and managers 3- During any other access to the platform
Dataset Date/Version	NA
Dataset Language	English
Related WP(s) and task(s)	WP2 - EMPATIA Platform Prototype, in T2.3 - Data analysis and information visualisation and T2.4 - Development and integration )
Data responsibilities	
Data Controller	Each Municipality/Entity where the platform is deployed
Data Processor	OneSource Consultoria Informática Lda
Other Parties with Access to Personal Data	OneSource Consultoria Informática Lda
Responsible for Data Storage	OneSource Consultoria Informática Lda
Data Collected and Standards	
List of Data Collected	Log Data: Component; User; IP address; Message; URL information; Date of creation; Date of update CPU Performance Data: CPU usage ratio; CPU idle ratio; CPU wait ratio Memory Performance Data: Memory usage ratio; Memory free ratio; Memory buffered ratio; Memory cached ratio; Memory swap usage ratio; Total physical memory I/O Performance Data: Number of IP packets received; Number of IP packets transmitted; Total bytes received; Total bytes transmitted; Disk read operations; Disk write operations; Web Performance Data: Number of requests per second; Number of connections
Standards (formats, estimated volume of data)	Formats/Standards: Storage: database, SQL for CRUD operations; Export of data: CSV, TXT, JSON or Excel file; Interoperability: JSON Estimated Volume: Dependent on the number of operations (from Kilobytes to Gigabytes)

Data Sharing	
Data Shared	Partially
Data sharing purpose (for dissemination or exploitation)	NA
Format of data sharing	NA
Means of data sharing	Published in D2.4 Development and integration of the platform – final (available on <a href="https://www.empatia-project.eu/">https://www.empatia-project.eu/</a> )
Embargo Period	NA
Open Data license	NA
Archiving and preservation	
Data Storage (including backups)	The dataset is preserved in the datacentre of OneSource, or where the platform is deployed.
Preservation periods and where	Data is kept for the duration of pilots and months after the conclusion date for analysis and platform enhancements.
Data Security	
Physical protection mechanisms	The datacenter where the data is archived and preserved has strict security policies regarding the physical access. The access to the datacenter is restricted to the CTO and CISO and its access is managed through two-level of security: physical key and alarm. All the access is logged in the alarm system and documented in the internal procedures of OneSource. The infrastructure for data archiving is supported by redundant servers and professional storage systems with active mechanisms of redundancy, and protection at the physical level for power supply (UPS systems). Storage uses fiber channel SAN with RAID and multiple servers. The backup system includes automated procedures, to assure data protection at two levels: In-datacenter backups and external-datacenter backups. The in-datacenter backups include daily copies automated and keeping an historic of two years. The external-datacenter backups are performed in a weekly basis and keep an history of two years. All the levels of backup employ encryption technologies and all the access to the backups is controlled in the SIEM of OneSource. Indeed, alerts of level 1 (i.e. marked with high severity) in the SIEM are scaled to the CISO for analysis regarding the possible security threats.
Logical protection mechanisms	The platform implements security mechanisms to protect the data and to manage the access to the data. As the platform is composed by several and independent components, these are distributed in different servers. All the data, especially personal data, is protected through specific components, that only allow access to the data to other components with valid JWT tokens (generated through valid authentication mechanisms). All the tokens have strict policies for expiration (i.e. 10 minutes), requiring new logins to provide valid tokens. As stated, all the transactions that require access to data require valid tokens. As such, no access to data can be performed without a valid login and access permissions, which are set according to the role of

	the user (e.g. if manager of an entity, or simply as an user participating in the PB process). All the accesses are analyzed in the SIEM of OneSource, where level 1 events (i.e. marked with high severity) are scaled to the CISO, while level 2 (i.e. marked with medium severity) and level 3 (i.e. marked with low severity) are managed by the network administrator and systems operators of OneSource.
Ethical remarks	
Legal framework	National Framework: “Artigo 35º da Constituição da República Portuguesa – utilização da informática; Lei 32/2008 - transpõe a Diretiva da Retenção de Dados, relativa à conservação de dados das comunicações eletrónicas”
Policies /Information Sheet	NA
Ethical risks	None

### 3.4. Dataset #4 – Scientific Dissemination

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/dataset-4-%E2%80%93-scientific-dissemination">https://dkan.empatia-project.online/dataset/dataset-4-%E2%80%93-scientific-dissemination</a>
Dataset Description	This dataset includes the information regarding the Scientific Dissemination of the project, including Scientific Articles submitted and approved for publications, books and chapters in books, papers presented to scientific conferences, articles in specialized magazines, and any other scientific result disseminated through the scientific community.
Dataset Purpose/Purposes	Public dissemination of scientific results of EMPATIA in compliance with the guidelines for Open Access provided by the EUROPEAN COMMISSION Directorate-General for Research & Innovation
Dataset Source	Data generated for scientific dissemination purpose under the activity of EMPATIA
Dataset Date/Version	Version 1.0 31/12/2016
Dataset Language	English
Related WP(s) and task(s)	WP5 - Dissemination and Exploitation
Data responsibilities	
Data Processor	Centro de Estudos Sociais
Data Controller	Centro de Estudos Sociais
Responsible for Data Storage	OneSource Consultoria Informática Lda
Data Collected and Standards	
List of Data Collected	<p>Tentative list of Metadata:</p> <ul style="list-style-type: none"> <li>- Abstract Short description of the work</li> <li>- Accessed Date an electronic resource was accessed. Typically filled automatically. Accepts various date formats, including “today,” “yesterday,” and “tomorrow”</li> <li>- Date Date of publication. See “Accessed” for data entry</li> <li>- DOI The Digital Object Identifier of an item. This field will be added to more item types in the future]]</li> <li>- Extra Free field. Currently used for storing PMID and PMCID. Note these have to be on separate lines to be usable in citations.</li> <li>- Format The format of an audio or video recording (e.g. “DVD,” “CD,” “MP3,” etc.)</li> <li>- ISBN The International Standard Book Number of a book or similar publication</li> <li>- ISSN The International Standard Serial Number of a periodical publication</li> <li>- Issue The issue number of a periodical - typically issues start at 1 at the beginning of each year/volume</li> </ul>

	<ul style="list-style-type: none"> <li>- Language The language of publication for an item. We recommend storing these as two letter ISO language codes followed by two letter ISO country codes, e.g. en-US for American English, or de-DE for German. Note that this also serves to disable title-casing of non-English publications.</li> <li>- Pages The page range of an item that is part of a larger publication. Also to be used for locators in e-journals</li> <li>- Place Mostly place of publication for an item. Can also refer to location of an event. Zotero will eventually add separate place fields for those two cases.</li> <li>- Publication The title of the periodical containing the cited item</li> <li>- Publisher The publisher of an item</li> <li>- License The copyright terms or license for an item</li> <li>- Section Section of a bill, statute, or a newspaper</li> <li>- Series Name of a series that contains multiple publications (e.g. “Cambridge Studies in Comparative Politics”)</li> <li>- Series Number The number of an item in a series</li> <li>- Series Title Title of a series of articles within one issue of a journal. See here for an explanation. For citation purposes this is currently equivalent to “Series” and is erroneously used instead of series in some item types (e.g. Map)</li> <li>- Title The principal title of an item. Should be entered in sentence case</li> <li>- Type Description of an item, such as “Ph.D. Thesis,” “Unpublished manuscript,” or “Poster”</li> <li>- URL URL (web-address) at which the full item was accessed. Should not be used for links to catalogue records or abstracts, which can be added as links</li> <li>- Volume volume of a (typically multi-volume) publication. Will most commonly be a number</li> </ul>
Standards (formats, estimated volume of data)	<p>Formats/Standards:  Storage: database, SQL for CRUD operations  Export of data: CSV, TXT, PDF or Excel File  Interoperability: JSON, PDF  Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)</p>
<b>Data Sharing</b>	
Data Shared	All data collected are shared
Data sharing purpose (for dissemination or exploitation)	To allow independent research and independent monitoring of the results of EMPATIA
Format of data sharing	CSV, TXT, DOC, ODT, PDF or Excel File
Means of data sharing	Data will be shared on on <a href="https://www.empatia-project.eu/">https://www.empatia-project.eu/</a> and other public repositories of scientific articles.
Embargo Period	Max 12 months
Open Data license	Self-archiving / 'green' open access on the repository hosted on <a href="http://empatia-project.eu/">http://empatia-project.eu/</a> .– The author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online repository

	<p>before, at the same time as, or after publication. Some publishers request that open access be granted only after an embargo period has elapsed.</p> <p>Open access publishing / 'gold' open access – an article is immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers. The most common business model is based on one-off payments by authors. In other cases, the costs of open access publishing are covered by subsidies or other funding models. Finally, a number of digital Scientific Journals directly adopt an Open Access and are published online free of charge.</p>
<b>Archiving and preservation</b>	
Data Storage (including backups)	NA
Preservation periods and where	NA
<b>Data Security</b>	
Physical protection mechanisms	NA
Logical protection mechanisms	NA
<b>Ethical remarks</b>	
Legal framework	NA
Policies /Information Sheet	NA
Ethical risks	NA

### 3.5. Dataset #5 – Dissemination

Dataset Identification	
Dataset Description	This dataset includes the information and personal data collected through the main website of the project <a href="http://empatia-project.eu/">http://empatia-project.eu/</a> . The main website will be used to disseminate the result of EMPATIA to the broader public possible and at the same time to feed and maintain a community of users and followers of EMPATIA.
Dataset Purpose/Purposes	For communication purposes, to maintain a constant flow of communication with the community of EMPATIA's users and followers; To stimulate public debate online regarding challenges and opportunities for Digital Democratic Innovations; To test in protected environment the new tools released by EMPATIA before their actual use in pilots.
Dataset Source	Data collected through the web application service based on EMPATIA and hosted on <a href="http://empatia-project.eu/">http://empatia-project.eu/</a> at the registration and authentication of new users.
Dataset Date/Version	Version 2.0 31/12/2017
Dataset Language	English
Related WP(s) and task(s)	WP5 - Dissemination and Exploitation
Data responsibilities	
Data Processor	Centro de Estudos Sociais
Data Controller	Centro de Estudos Sociais
Responsible for Data Storage	OneSource Consultoria Informática Lda
Data Collected and Standards	
List of Data Collected	<p>a) Personal Data:</p> <ul style="list-style-type: none"> <li>Name Surname</li> <li>Email</li> <li>Date of birth</li> <li>Neighborhood/Address(optional)</li> <li>Telephone   Mobile phone (optional)</li> <li>NIF (optional)</li> <li>Genre (optional)</li> <li>Profession (optional)</li> <li>Education (optional)</li> </ul> <p>c) Non-personally identifiable information</p>

	<p>Browser type  Language preference  Referring site  Date and time of each visitor request</p> <p>d) Potentially personally-identifiable information (see also Dataset #4)  IP address</p>
Standards (formats, estimated volume of data)	<p>Formats/Standards:  Storage: database, SQL for CRUD operations  Export of data: CSV, TXT, PDF or Excel File  Interoperability: JSON, PDF  Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)</p>
<b>Data Sharing</b>	
Data Shared	NA
Data sharing purpose (for dissemination or exploitation)	NA
Format of data sharing	NA
Means of data sharing	NA
Embargo Period	NA
Open Data license	NA
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved in the OneSource datacentre. There are daily backups and weekly backups to a dedicated server.
Preservation periods and where	The collected data in this dataset is kept no longer than one year after the last access to the Service by the Data Collector. The data is kept in the datacenter(s) of OneSource.
<b>Data Security</b>	
Physical protection mechanisms	The datacenter where the data is archived and preserved has strict security policies regarding the physical access. The access to the datacenter is restricted to the CTO and CISO and its access is managed through two-level of security: physical key and alarm. All the access is logged in the alarm system and documented in the internal procedures of OneSource. The infrastructure for data archiving is supported by redundant servers and professional storage systems with active mechanisms of redundancy, and protection at the physical level for power supply (UPS systems). Storage

	<p>uses fiber channel SAN with RAID and multiple servers. The backup system includes automated procedures, to assure data protection at two levels: In-datacenter backups and external-datacenter backups. The in-datacenter backups include daily copies automated and keeping an historic of two years. The external-datacenter backups are performed in a weekly basis and keep an history of two years. All the levels of backup employ encryption technologies and all the access to the backups is controlled in the SIEM of OneSource. Indeed, alerts of level 1 (i.e. marked with high severity) in the SIEM are scaled to the CISO for analysis regarding the possible security threats.</p>
Logical protection mechanisms	<p>The platform implements security mechanisms to protect the data and to manage the access to the data. As the platform is composed by several and independent components, these are distributed in different servers. All the data, especially personal data, is protected through specific components, that only allow access to the data to other components with valid JWT tokens (generated through valid authentication mechanisms). All the tokens have strict policies for expiration (i.e. 10 minutes), requiring new logins to provide valid tokens. As stated, all the transactions that require access to data require valid tokens. As such, no access to data can be performed without a valid login and access permissions, which are set according to the role of the user (e.g. if manager of an entity, or simply as an user participating in the PB process). All the accesses are analyzed in the SIEM of OneSource, where level 1 events (i.e. marked with high severity) are scaled to the CISO, while level 2 (i.e. marked with medium severity) and level 3 (i.e. marked with low severity) are managed by the network administrator and systems operators of OneSource.</p>
Ethical remarks	
Legal framework	<p>International Framework: Directive 95/46/EC . National Framework: "Artigo 35º da Constituição da República Portuguesa – utilização da informática Lei 67/ 98 – Lei da proteção de Dados Pessoais Lei 41/2004 - Regula a proteção de dados pessoais no sector das Comunicações Eletrónicas (alterada e republicada) Lei 32/2008 - transpõe a Diretiva da Retenção de Dados, relativa à conservação de dados das comunicações eletrónicas "</p>
Policies /Information Sheet	<p>Privacy Policy: <a href="https://www.empatia-project.eu/page/privacy_policy">https://www.empatia-project.eu/page/privacy_policy</a> Terms of Service: <a href="https://www.empatia-project.eu/page/use_terms">https://www.empatia-project.eu/page/use_terms</a></p>
Ethical risks	<p>Dataset includes Personal and Sensitive Data. Adequate security measures and protocols shall be implemented by the Data Controller in order to ensure the enforcement of the protection of personal data according to the provisions detailed in the privacy policies.</p>

### 3.6. Dataset #6 – Mapping Participatory Innovations Survey

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/dataset-6-%E2%80%93-mapping-participatory-innovations-survey">https://dkan.empatia-project.online/dataset/dataset-6-%E2%80%93-mapping-participatory-innovations-survey</a>
Dataset Description	This dataset includes the information and personal data collected through the survey website of the project <a href="https://oidp.empatia-project.eu/">https://oidp.empatia-project.eu/</a> . The survey website is sponsored by the OIDP network, the OIDP 2017 conference, the EMPATIA project, and Participedia project. The objective of this survey is to collect and share data about a variety of participatory processes implemented by OIDP partners, and at the same time to feed and maintain a community of users and followers of democratic innovations which are related to EMPATIA.
Dataset Purpose	<p>For mapping participatory innovations around the world;</p> <p>For creating a network of people and organisations interested in democratic innovations, to share and learn about best initiatives/approaches;</p> <p>To stimulate public debate online regarding challenges and opportunities for Digital Democratic Innovations;</p> <p>To communicate about EMPATIA through the network.</p> <p> Dataset Source   Data collected through the web application service based on EMPATIA and hosted on <a href="https://oidp.empatia-project.eu/">https://oidp.empatia-project.eu/</a>.</p> <p>At the end of the survey, users can voluntarily provide their name and email to be part of OIDP multichannel working group.</p> <p>During the submission of survey answers, non-attributable</p> <p> Dataset Date/Version   Version 1.0 31/03/2017 </p>
Dataset Language	English, French, Portuguese and Spanish
Related WP(s) and task(s)	WP5 - Dissemination and Exploitation
Data responsibilities	
Data Processor	Centro de Estudos Sociais
Data Controller	Centro de Estudos Sociais
Responsible for Data Storage	OneSource Consultoria Informática Lda
Data Collected and Standards	
List of Data Collected	<p>a) Personal Data:</p> <ul style="list-style-type: none"> <li>- Name Surname</li> <li>- Email</li> </ul> <p>b) Survey multiple-choice answers and comments</p> <p>c) Non-personally identifiable information</p> <ul style="list-style-type: none"> <li>- Browser type</li> <li>- Language preference</li> </ul>

	<ul style="list-style-type: none"> <li>- Referring site</li> <li>- Date and time of each visitor request</li> </ul> <p>d) Potentially personally-identifiable information (see also Dataset #4)</p> <ul style="list-style-type: none"> <li>- IP address</li> </ul>
Standards (formats, estimated volume of data)	<p>Formats/Standards:  Storage: database, SQL for CRUD operations  Export of data: CSV, TXT, PDF or Excel File  Interoperability: JSON, PDF  Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)</p>
<b>Data Sharing</b>	
Data Shared	a) Anonymized results of the Survey
Data sharing purpose (for dissemination or exploitation)	To allow independent research of survey results
Format of data sharing	
Means of data sharing	On the website <a href="https://oidp.empatia-project.eu/">https://oidp.empatia-project.eu/</a> , and on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>
Embargo Period	Max 12 months
Open Data license	Depends by the mean of data sharing
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved in the OneSource datacentre. There are daily backups and weekly backups to a dedicated server.
Preservation periods and where	The collected data in this dataset is kept no longer than one year after the last access to the Service by the Data Collector. The data is kept in the datacenter(s) of OneSource.
<b>Data Security</b>	
Physical protection mechanisms	The datacenter where the data is archived and preserved has strict security policies regarding the physical access. The access to the datacenter is restricted to the CTO and CISO and its access is managed through two-level of security: physical key and alarm. All the access is logged in the alarm system and documented in the internal procedures of OneSource. The infrastructure for data archiving is supported by redundant servers and professional storage systems with active mechanisms of redundancy, and protection at the physical level for power supply (UPS systems). Storage uses fiber channel SAN with RAID and multiple servers. The backup system includes automated procedures, to assure data protection at two levels: In-datacenter backups and external-datacenter backups. The in-datacenter backups include daily copies automated and keeping an historic of two years. The external-datacenter backups are performed in a weekly basis and keep an history of two years. All the levels of backup employ encryption technologies and all the access to the backups is controlled in the

	SIEM of OneSource. Indeed, alerts of level 1 (i.e. marked with high severity) in the SIEM are scaled to the CISO for analysis regarding the possible security threats.
Logical protection mechanisms	The platform implements security mechanisms to protect the data and to manage the access to the data. As the platform is composed by several and independent components, these are distributed in different servers. All the data, especially personal data, is protected through specific components, that only allow access to the data to other components with valid JWT tokens (generated through valid authentication mechanisms). All the tokens have strict policies for expiration (i.e. 10 minutes), requiring new logins to provide valid tokens. As stated, all the transactions that require access to data require valid tokens. As such, no access to data can be performed without a valid login and access permissions, which are set according to the role of the user (e.g. if manager of an entity, or simply as an user participating in the PB process). All the accesses are analyzed in the SIEM of OneSource, where level 1 events (i.e. marked with high severity) are scaled to the CISO, while level 2 (i.e. marked with medium severity) and level 3 (i.e. marked with low severity) are managed by the network administrator and systems operators of OneSource.
Ethical remarks	
Legal framework	Directive 95/46/EC
Policies /Information Sheet	Privacy Policy: <a href="https://oidp.empatia-project.eu/auth/privacyPolicy">https://oidp.empatia-project.eu/auth/privacyPolicy</a> Terms of Service: <a href="https://oidp.empatia-project.eu/auth/useTerms">https://oidp.empatia-project.eu/auth/useTerms</a>
Ethical risks	Dataset could include Personal Data. Adequate security measures and protocols shall be implemented by the Data Controller in order to ensure the enforcement of the protection of personal data according to the provisions detailed in the privacy policies.

### 3.7. Dataset #7 – Evaluation & Impact Assessment

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/dataset-7-%E2%80%93-evaluation-impact-assessment">https://dkan.empatia-project.online/dataset/dataset-7-%E2%80%93-evaluation-impact-assessment</a>
Dataset Description	<p>This dataset includes the data collected for the purposes of the evaluation and impact assessment of the project. The evaluation of the EMPATIA pilots, platform and approach has a dual focus: on one hand it evaluates the process perspective of PB to assess its impact on e-participation, transparency and democratic process, and on the other hand it evaluates the technical, user related and behavioural aspects.</p> <p>While the data controller for this Dataset is Brunel University, almost all partners involved in the activity of WP3 have been directly engaged in processing data collected online and in person settings.</p>
Dataset Purpose	<ul style="list-style-type: none"> <li>- To evaluate the impact of EMPATIA on the social, economic, behavioural and political dimensions of pilots</li> <li>- To evaluate the operational performance (technical and behavioural) of the EMPATIA platform, when used in three different target communities.</li> <li>- To derive feedback from the communities; provide an overall impact assessment of the project results for internal input and for future improvements.</li> </ul>
Dataset Source	<p>Data collected through:</p> <ol style="list-style-type: none"> <li>1- Online facultative questionnaires through the websites of pilots               <ol style="list-style-type: none"> <li>a. In Lisbon based on the EMPATIA platform (see also Dataset #8 – Pilot - Lisbon)</li> <li>b. In Ričany based on the ridimričany platform (see also Dataset #10 – Pilot – Říčany)</li> </ol> </li> <li>2- Online facultative questionnaires delivered through a third party service (<a href="https://surveyMonkey.com/">https://surveyMonkey.com/</a>) managed by CES.</li> <li>3- Facultative interviews and focus groups with the participants to the pilots (All pilots)</li> </ol>
Dataset Date/Version	Foreseen (March 2018)
Dataset Language	English
Related WP(s) and task(s)	WP4 - Evaluation and Impact Assessment
Data responsibilities	
Data Processor	Brunel University London, Centro de Estudos Sociais, OneSource, D21, ZebraLog, UNIMI, InLOCO
Data Controller	Brunel University London
Responsible for Data Storage	Brunel University London, Centro de Estudos Sociais, OneSource, D21
Data Collected and Standards	
List of Data Collected	KPIs

- Technical KPIS
- Behavioural KPIS
- Socio-economic KPIS
- Political KPIS: inclusiveness and political alienation

KPIs category	KPIs	Evaluation method
Performance Expectancy	Perceived Usefulness	Survey after field trial
	Extrinsic Motivation	Survey after field trial
	Job-fit	Survey after field trial
	Relative Advantage	Survey after field trial
	Outcome Expectations	Survey after field trial
Effort Expectancy	Perceived Ease of Use	Survey after field trial
	Complexity	Survey after field trial
	Ease of Use	Survey after field trial
Social Influence	Subjective Norm	Survey after field trial
	Social Factors	Survey after field trial
	Image	Survey after field trial
Facilitating Conditions	Perceived Behavioural Control	Survey after field trial
	Facilitating Conditions	Survey after field trial
	Compatibility	Survey after field trial
KPIs category	KPIs	Evaluation method
System Quality	Reliability	Survey after field trial
	Flexibility	Survey after field trial
	Integration	Survey after field trial
	Accessibility	Survey after field trial
	Timeliness	Survey after field trial
Information Quality	Completeness	Survey after field trial
	Accuracy	Survey after field trial
Service Quality	Tangibles	Survey after field trial
	Reliability	Survey after field trial
	Responsiveness	Survey after field trial
	Assurance	Survey after field trial
	Empathy	Survey after field trial
Information Use	Usefulness	Survey after field trial
	Ease of Use	Survey after field trial
User Satisfaction	System Satisfaction	Survey after field trial
Willingness to provide personal information to the e-service	Perceived Internet privacy risk	Survey after field trial
	Internet privacy concerns	Survey after field trial

	Internet trust	Survey after field trial
	Personal Internet interest	Survey after field trial

#### Summary of Generic Socio-Economic KPIs for EMPATIA

KPIs category	KPIs	Evaluation method
Cost Saving	Money saving	Survey after field trial
	Time Saving	Survey after field trial
Openness	Openness	Survey after field trial
Trust	Trust in the Internet	Survey after field trial
	Trust in the organisation	Survey after field trial
Business/Operational Issues	Operational cost	Survey after field trial
	Capital expenditure	Survey after field trial
	Cost of migration	Survey after field trial
	Vendor lock-in	Survey after field trial
Legal and regulatory compliance	Forensics	Survey after field trial
	Data retention and track back	Survey after field trial
	Organisation's control over the data	Survey after field trial

#### Political-KPIs for Assessing EMPATIA Platform

KPIs category	Description	Objectives
Inclusiveness I: who participates?	Socio Demographic Data; Progressive/Conservative attitudes, Attention to Politics, past voting behaviour	Explore the profile of the participants
Inclusiveness II: Channel elasticity	Metrics that evaluates the willingness of the participants to change channel from online to face-to-face and vice versa	Explore substitution effects of the presence of multiple channels.
Political Alienation I: Efficacy	Internal & external efficacy	Explore the impact of participating in the process on efficacy indicators

	Political Alienation II: Anti-politics	Systemic and local measures of trust and anti-politics	Explore the impact of participating in the process on trust and anti-politics indicators
Standards (formats, estimated volume of data)	Formats/Standards: CSV, ODT, DOC, images (GIF). Estimated Volume: Below 200 Mbytes.		
<b>Data Sharing</b>			
Data Shared	All data collected will be shared after undergoing a proper anonymization process. No personal data will be shared		
Data sharing purpose (for dissemination or exploitation)	For transparency purposes and to improve the quality of the Social, Ethical and Legal analysis of the pilot of EMPATIA.		
Format of data sharing	Data will be released: - in machine readable format (CSV) - as charts and other info-graphic elaborations (PDF) - images (GIF)		
Means of data sharing	Published in D4.2 - Evaluation and Pilots impact assessment, and on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>		
Embargo Period	NA		
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>		
<b>Archiving and preservation</b>			
Data Storage (including backups)	<p>Along the data collection, data have been temporarily stored in different manner, according to the channel of collection:</p> <ol style="list-style-type: none"> <li>1- Online questionnaires through the websites of pilots <ol style="list-style-type: none"> <li>a. In Lisbon, stored on the Server of OneSource, where is installed the related instance of the EMPATIA platform (see also Dataset #8 – Pilot - Lisbon)</li> <li>b. In Riçany based on the ridimriçany platform (see also Dataset #10 – Pilot – Řiçany)</li> </ol> </li> <li>2- Online questionnaires delivered through a third party service (<a href="https://surveymonkey.com/">https://surveymonkey.com/</a>) managed by CES have been kept on the server provided by Survey Monkey. The specific conditions can be read here: <a href="https://www.surveymonkey.com/mp/policy/privacy-policy/">https://www.surveymonkey.com/mp/policy/privacy-policy/</a> . (All Pilots)</li> <li>3- Data collected in interviews and focus groups with the participants to the pilots (All Pilots) have been archived directly by the data controller in private files hosted in the personal computer owned by Brunel University.</li> </ol> <p>During the processing operation, personal identifiable information have been pseudonymized and resulting data have been kept in remote files (various formats) hosted in the personal computer owned by Brunel University and Center for Social Studies of the University of Coimbra.</p>		

Preservation periods and where	Personal data will be destroyed at the conclusion date of EMPATIA.
<b>Data Security</b>	
Physical protection mechanisms	NA
Logical protection mechanisms	NA
<b>Ethical remarks</b>	
Legal framework	University Research Ethics Committee of Brunel University approved the research (letter annexed to D4.1 - Evaluation plans and guidelines)
Policies /Information Sheet	Published in D4.1 - Evaluation plans and guidelines, and D1.5 – Social, Ethical and Legal analysis, and presented to the individuals involved, prior to data collection.
Ethical risks	<p>Medium Risk</p> <p>Dataset includes Personal and Sensitive Data and involves a large number of data processor, in charge of data collection in four different European countries, under an unique coordination from the data controller.</p> <p>The peculiar structure of this dataset reflects the complexity of the partnership of EMPATIA and the transnational transnational scope of its research activity.</p> <p>Specific risks are detected regarding the use of third party software to deliver part of the questionnaires.</p> <p>Adequate security measures and protocols shall be implemented by the Data Controller in order to ensure the enforcement of the protection of personal data according to the provisions detailed in the privacy policy.</p>

### 3.8. Dataset #8 – Pilot - Lisbon

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/empatia-db-8-pilot-lisbon">https://dkan.empatia-project.online/dataset/empatia-db-8-pilot-lisbon</a>
Dataset Description	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Lisbon (Portugal). The pilot focuses on the integration between different services that correspond to various Democratic innovations experimented in Lisbon as the Participatory Budgeting, Public Debates, Fix my Street and Open Data sharing. The EMPATIA's service should provide the framework for the integration of pre-existing digital services (managed by same municipality of Lisbon as well as by independent third parties) in the site: <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a>
Dataset PurposePurposes	<ul style="list-style-type: none"> <li>- To ensure the unique authentication of users, necessary to take part to processes managed on <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a>;</li> <li>- To study and research how users and visitors use <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a>;</li> <li>- To communicate with users regarding eventual updates to <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a>, to EMPATIA's services and to related policies;</li> <li>- To provide periodical information regarding the content of <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a>, in accordance with the notification preferences configured by each user;</li> </ul>
Dataset Source	Data collected through the web application service based on EMPATIA and hosted on <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a> : <ul style="list-style-type: none"> <li>- At the registration and authentication of new users</li> <li>- During the submission of comments, votes and generation of any other user content</li> <li>- During any other access to the platform</li> </ul>
Dataset Date/Version	1.0 December 2017
Dataset Language	Portuguese
Related WP(s) and task(s)	WP1 - Foundations and Models for Multichannel Participatory Budgeting; T1.1 - Theories, models and cases studies for Participatory Budgeting; T1.3 - Social, ethics and legal analysis; WP3 - Pilots; T3.2 - Pilot: Lisbon/Portugal
Data responsibilities	
Data Processor	OneSource Consultoria Informática Lda
Data Controller	OneSource Consultoria Informática Lda
Responsible for Data Storage	OneSource Consultoria Informática Lda
Data Collected and Standards	
List of Data Collected	a) Personal Data: Name Surname

	<p>Email  Age range  Neighborhood/Address  Telephone   Mobile phone (optional)  Gender  Education  b) Votes and preferences.  c) Non-personally identifiable information  Browser type  Language preference  Referring site  Date and time of each visitor request  d) Potentially personally-identifiable information (see also Dataset #4)  IP address  e) User Content and Personally Identifiable Information:  Proposals related content and comments to proposals  Content produced in online debates</p>
Standards (formats, estimated volume of data)	<p>Formats/Standards:  Storage: database, SQL for CRUD operations  Export of data: CSV, TXT, PDF or Excel File  Interoperability: JSON, PDF  Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)</p>
<b>Data Sharing</b>	
Data Shared	<p>a) Anonymized and clustered Personal Data:  Age  Neighborhood/Address  Gender  Education  b) Anonymized results of voting/preference session  c) User Content and Personally Identifiable Information:  Proposals related content  Any other public user generated content</p>
Data sharing purpose (for dissemination or exploitation)	To allow independent research and independent monitoring of the delivery and outcomes of the Pilot of Lisbon in compliance with the Open Data principles established for EMPATIA.
Format of data sharing	Data will be released: in machine readable format (CSV)

	as charts and other info-graphic elaborations (PDF) images (GIF)
Means of data sharing	On the website <a href="http://www.lisboaparticipa.pt">http://www.lisboaparticipa.pt</a> and on the repository hosted on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>
Embargo Period	6 months
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved in the OneSource datacentre. There are daily backups and weekly backups to a dedicated server.
Preservation periods and where	The collected data in this dataset is kept no longer than one year after the last access to the Service by the Data Collector. The data is kept in the datacenter(s) of OneSource.
<b>Data Security</b>	
Physical protection mechanisms	The datacenter where the data is archived and preserved has strict security policies regarding the physical access. The access to the datacenter is restricted to the CTO and CISO and its access is managed through two-level of security: physical key and alarm. All the access is logged in the alarm system and documented in the internal procedures of OneSource. The infrastructure for data archiving is supported by redundant servers and professional storage systems with active mechanisms of redundancy, and protection at the physical level for power supply (UPS systems). Storage uses fiber channel SAN with RAID and multiple servers. The backup system includes automated procedures, to assure data protection at two levels: In-datacenter backups and external-datacenter backups. The in-datacenter backups include daily copies automated and keeping an historic of two years. The external-datacenter backups are performed in a weekly basis and keep an history of two years. All the levels of backup employ encryption technologies and all the access to the backups is controlled in the SIEM of OneSource. Indeed, alerts of level 1 (i.e. marked with high severity) in the SIEM are scaled to the CISO for analysis regarding the possible security threats.
Logical protection mechanisms	The platform implements security mechanisms to protect the data and to manage the access to the data. As the platform is composed by several and independent components, these are distributed in different servers. All the data, especially personal data, is protected through specific components, that only allow access to the data to other components with valid JWT tokens (generated through valid authentication mechanisms). All the tokens have strict policies for expiration (i.e. 10 minutes), requiring new logins to provide valid tokens. As stated, all the transactions that require access to data require valid tokens. As such, no access to data can be performed without a valid login and access permissions, which are set according to the role of the user (e.g. if manager of an entity, or simply as an user participating in the PB process). All the accesses are analyzed in the SIEM of OneSource, where level 1 events (i.e. marked with high severity) are scaled to the CISO, while level 2 (i.e. marked with medium severity) and level 3 (i.e. marked with low severity) are managed by the network administrator and systems operators of OneSource.
<b>Ethical remarks</b>	

Legal framework	<p>International Framework:          Directive 95/46/EC .          National Framework:          “Artigo 35º da Constituição da República Portuguesa – utilização da informática          Lei 67/ 98 – Lei da proteção de Dados Pessoais          Lei 41/2004 - Regula a proteção de dados pessoais no sector das Comunicações Eletrónicas (alterada e republicada)          Lei 32/2008 - transpõe a Diretiva da Retenção de Dados, relativa à conservação de dados das comunicações eletrónicas          ”          National Regulatory Body: CNPD (<a href="https://www.cnpd.pt/">https://www.cnpd.pt/</a>)</p>
Policies /Information Sheet	NA
Ethical risks	<p>Dataset includes Personal and Sensitive Data. Adequate security measures and protocols shall be implemented by the Data Controller in order to ensure the enforcement of the protection of personal data according to the provisions detailed in the privacy policies.</p>

### 3.9. Dataset #9 – Pilot - Milan

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/empatia-db-9-pilot-milan">https://dkan.empatia-project.online/dataset/empatia-db-9-pilot-milan</a>
Dataset Description	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Milan (Italy) through the dedicated website <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a> . The pilot focuses on both the two cycles of the Participatory Budgeting process and their interaction: the implementation of the winning projects of the first PB edition and the deliberation cycle of the second PB edition. The pilot uses two different platforms, one for each cycle. Data from the Deliberation Cycle refers only to the first phase.
Dataset Purposes	<p>Purposes</p> <p>To ensure the unique authentication of users, necessary to take part to processes managed on <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>;</p> <p>To study and research how users and visitors use <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>;</p> <p>To communicate with users regarding eventual updates to <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>, to EMPATIA's services and to related policies;</p> <p>To provide periodical information regarding the content of <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>, in accordance with the notification preferences configured by each user.</p>
Dataset Source	<p>Data collected through the web application service based on OPENDCN (<a href="http://www.opendcn.org">http://www.opendcn.org</a>) and hosted on <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>:</p> <ul style="list-style-type: none"> <li>- At the registration and authentication of new users</li> <li>- During the submission of comments, votes and generation of any other user content</li> <li>- During any other access to the platform.</li> </ul> <p>Data collected through the web application service based on EMPATIA and hosted on <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>: Limited to User Generated Content (Public)</p>
Dataset Date/Version	Foreseen
Dataset Language	Italian
Related WP(s) and task(s)	WP1 - Foundations and Models for Multichannel Participatory Budgeting; T1.1 - Theories, models and cases studies for Participatory Budgeting; T1.3 - Social, ethics and legal analysis; WP3 - Pilots; T3.4 - Pilot: Milan/Italy
Data responsibilities	
Data Processor	Università degli Studi di Milano
Data Controller	Comune di Milano
Responsible for Data Storage	Università degli Studi di Milano

## Data Collected and Standards

List of Data Collected	a) Personal Data: Name, Surname Email Status (Resident/City User/nd) Age Gender Education Neighborhood (only for the residents in the advanced registration) Individual tax code (only in the advanced registration) Mobile phone (only for the residents in the advanced registration) A document (pdf, image, ...) for attesting that user is a city user (only in the advanced registration) b) Votes and Preferences c) Non-personally identifiable information Browser type Language preference Referring site Date and time of each visitor request d) Potentially personally-identifiable information (see also Dataset #4) IP address e) User Content and Personally Identifiable Information: Proposals related content and comments to proposals Content produced in online debates .
Standards (formats, estimated volume of data)	Formats/Standards: Storage: database, SQL for CRUD operations Export of data: CSV, TXT, PDF or Excel File Interoperability: JSON, PDF Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)
<h2>Data Sharing</h2>	
Data Shared	a) Anonymized and clusterized Personal Data: Age Neighborhood/Address Gender Education b) Anonymized votes and preferences e) User Content Proposals related content

	Any other public user generated content
Data sharing purpose (for dissemination or exploitation)	To allow independent research and independent monitoring of the delivery and outcomes of the Pilot of Milan in compliance with the Open Data principles established for EMPATIA.
Format of data sharing	Data will be released: in machine readable format (CSV) as charts and other info-graphic elaborations (PDF) images (GIF)
Means of data sharing	On the website <a href="https://www.bilanciopartecipativomilano.it">https://www.bilanciopartecipativomilano.it</a> and on the repository hosted on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>
Embargo Period	6 months
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved in the Università di Milano - Dipartimento di Informatica datacentre. There are daily backups to a dedicated server.
Preservation periods and where	The collected data in this dataset is kept no longer than one year after the last access to the Service by the Data Collector. The data is kept in the datacenter(s) of Università di Milano - Dipartimento di Informatica.
<b>Data Security</b>	
Physical protection mechanisms	The room in which the servers are located and where the data are stored has strict security rules regarding physical access. Access is limited to staff members of IT facilities of Università di Milano - Dipartimento di Informatica. The data storage infrastructure is supported by redundant servers and storage systems with active redundancy mechanisms covering both the disks and the connections between storage and servers that make up the infrastructure. The backup system includes automated procedures to perform data backup operations on a daily basis. Backups maintain a two-month history.
Logical protection mechanisms	In the logical protection many levels cooperate to preserve the integrity of the system: External level. The incoming/outcoming traffic is analyzed from Unimi firewall to prevent potential malicious requests OS level. The OS uses a firewall to block incoming request to not authorized services and uses the MAC system rules, using Selinux, to prevent a intrusion/manumission of the system. Web server level. The web server uses security module to mitagate potentials brute force attacks. Application level. The software implements a suited granularity of system permissions and the administrator can't modify directly the sensitive information of users to preserve integrity of the system.
<b>Ethical remarks</b>	
Legal framework	International Framework: Directive 95/46/EC

	<p>National Framework: Legge 31 dicembre 1996, n. 675 National Regulatory Body: Garante per la protezione dei dati personali (<a href="http://www.garanteprivacy.it/">http://www.garanteprivacy.it/</a>)</p>
Policies /Information Sheet	<p><a href="https://www.bilanciopartecipativomilano.it/policies/privacy">https://www.bilanciopartecipativomilano.it/policies/privacy</a></p>
Ethical risks	<p>The dataset is built combining two different sources, but all personal data have been collected and processed through the service OPENDCN (<a href="http://www.opendcn.org">http://www.opendcn.org</a>) and hosted on <a href="http://www.bilanciopartecipativomilano.it">http://www.bilanciopartecipativomilano.it</a>, while the service based on EMPATIA has been used only to publish public data and content. The privacy policy proposed by the consortium has been reviewed and accepted by the Data Controller (the Municipality of Milan).</p>

### 3.10. Dataset #10 – Pilot – Říčany

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/empatia-db-10-pilot-ri%C3%A7any">https://dkan.empatia-project.online/dataset/empatia-db-10-pilot-ri%C3%A7any</a>
Dataset Description	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Říčany (CZ Republic) through the dedicated website: <a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a> . The EMPATIA's service is deployed in this pilot in integration with the services of the partner D21. The Integration between different technologies is still under definition and will be reported in an update to the present dataset description before the start of the collection of Personal Data.
Dataset PurposePurposes	<p>Purposes</p> <ul style="list-style-type: none"> <li>- To ensure the unique authentication of users, necessary to take part to processes managed the website: <a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a></li> <li>- To study and research how users and visitors use<a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a>;</li> <li>- To communicate with users regarding eventual updates to<a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a>, to EMPATIA's services and to related policies;</li> <li>- To provide periodical information regarding the content of<a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a>, in accordance with the notification preferences configured by each user;.</li> </ul>
Dataset Source	<p>Data collected :</p> <ol style="list-style-type: none"> <li>a) Through the web application service based on EMPATIA and hosted on <a href="https://www.prekvaptericany.cz/">https://www.prekvaptericany.cz/</a> <ul style="list-style-type: none"> <li>o Limited to UCG, description of proposals</li> </ul> </li> <li>b) Through the web application service based on D21 proprietary Service and hosted on Ridim Ricany (<a href="http://www.ridimricany.cz/">http://www.ridimricany.cz/</a>) <ul style="list-style-type: none"> <li>o At the registration and authentication of new users</li> <li>o During the submission of comments, votes and generation of any other user content</li> <li>o Through the migration of data already pre-existing on ridim ričany before the start of the pilot</li> </ul> </li> </ol>
Dataset Date/Version	1.0 31/12/2017
Dataset Language	Czech
Related WP(s) and task(s)	WP3 - Pilots;T3.3 - Pilot: Říčany/Czech Republic
Data responsibilities	
Data Processor	Demokracie 2.1
Data Controller	Demokracie 2.1
Responsible for Data Storage	OneSource Consultoria Informática Lda; Demokracie 2.1

Data Collected and Standards	
List of Data Collected	a) Personal Data: Name Surname Email Gender Age range Telephone   Mobile phone (optional) b) Votes and preferences c) User Content and Personally Identifiable Information: Proposals related content and comments to proposals
Standards (formats, estimated volume of data)	Formats/Standards: Storage: database, SQL for CRUD operations Export of data: CSV, TXT, PDF or Excel File Interoperability: JSON, PDF Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)
Data Sharing	
Data Shared	a) Anonymized and clusterized Personal Data: Gender Age range b) Anonymized results of Votes and preferences c) User Content Proposals related content and comments to proposals Any other public user generated content
Data sharing purpose (for dissemination or exploitation)	To allow independent research and independent monitoring of the delivery and outcomes of the Pilot of Ricany in compliance with the Open Data principles established for EMPATIA.
Format of data sharing	Data will be released: in machine readable format (CSV) as charts and other info-graphic elaborations (PDF) images (GIF)
Means of data sharing	On the repository hosted on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>
Embargo Period	6 months
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>
Archiving and preservation	

Data Storage (including backups)	NA
Preservation periods and where	NA
<b>Data Security</b>	
Physical protection mechanisms	NA
Logical protection mechanisms	NA
<b>Ethical remarks</b>	
Legal framework	<p>International Framework: Directive 95/46/EC .</p> <p>National Framework: National Framework: The Office for Personal Data Protection is an independent body set up to supervise observance of the legal obligations laid down for processing of personal data; maintain the register of notified data processing operations; deal with initiatives and complaints from citizens concerning breach of law; and provide consultancy in personal data protection.</p> <p>Competence in the field of personal data protection: The Office's scope of authority is stipulated by the Data Protection Act (Articles 2 and 29). Moreover, the Office is vested with additional powers related to special issues and anchored in special laws.</p> <p>In developing its administrative activities, the Office acts on the Data Protection Act and the Administrative Code (Act No. 500/2004 Coll.) which is the basic piece of procedural law.</p> <p>In its supervisory activities, the Office is governed by the provisions of Act No. 101/2000 Coll. and certain provisions of Act No. 552/1991 Coll., on state control. The specimen of the identification card held by the Office's inspectors is depicted in the Government Order No. 8/2005 Coll.</p> <p>The Office administers the public administration information systems pursuant to Act No. 365/2000 Coll., on information systems of public administration.</p> <p>The Office provides information relating to its scope of authority pursuant to Act No. 106/1999 Coll., on free access to information.</p> <p>National Regulatory Body: Office for Personal Data (<a href="https://www.uoou.cz/en/">https://www.uoou.cz/en/</a>)</p>
Policies /Information Sheet	<a href="https://ricany.empatia-project.eu/page/privacy_policy">https://ricany.empatia-project.eu/page/privacy_policy</a> (On the EMPATIA-based site); No privacy policy available on Ridim Ricany ( <a href="http://www.ridimricany.cz/">http://www.ridimricany.cz/</a> )
Ethical risks	<p>The dataset is built combining two different sources, but all personal data have been collected and processed through the service Ridim Ricany (<a href="http://www.ridimricany.cz/">http://www.ridimricany.cz/</a>), managed by D21 and pre-existing the EMPATIA project, while the service based on EMPATIA has been used only to publish public data and content.</p> <p>The ethical coordinator cannot guarantee the respect of ethical standards for data collected on Ridim Ricany prior to the project. All data have been transmitted from the data controller to the partners already anonymized.</p>

### 3.11. Dataset #11 –Pilot – Wuppertal

Dataset Identification	
Repository	<a href="https://dkan.empatia-project.online/dataset/empatia-db-11-pilot-wuppertal">https://dkan.empatia-project.online/dataset/empatia-db-11-pilot-wuppertal</a>
Dataset Description	This dataset includes the information and personal data collected during the delivery of the EMPATIA's pilot in the municipality of Wuppertal (Germany) through the dedicated website: <a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a> . The EMPATIA's service is deployed in this pilot in integration with the services of the partner Zebralog.
Dataset PurposePurposes	Purposes <ul style="list-style-type: none"> <li>• To ensure the unique authentication of users, necessary to take part to processes managed the website: <a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a></li> <li>• To study and research how users and visitors use <a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a>;</li> <li>• To communicate with users regarding eventual updates to<a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a>, to EMPATIA's services and to related policies;</li> <li>• To provide periodical information regarding the content of<a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a>, in accordance with the notification preferences configured by each user;</li> </ul>
Dataset Source	Data collected through the web application service based on EMPATIA and hosted on <a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a> : <ul style="list-style-type: none"> <li>• At the registration and authentication of new users</li> <li>• During the submission of comments, votes and generation of any other user content</li> <li>• During any other access to the platform</li> </ul>
Dataset Date/Version	Foreseen
Dataset Language	German
Related WP(s) and task(s)	WP3 - Pilots; T3.1 - Pilot: Bonn/Germany
Data responsibilities	
Data Processor	Zebralog
Data Controller	Zebralog
Responsible for Data Storage	OneSource Consultoria Informática Lda
Data Collected and Standards	
List of Data Collected	a) Personal Data: Name Surname Email

	<p>Age range          Neighborhood/Address          Telephone   Mobile phone (optional)          Genre          Education          b) Votes and preferences          c) Non-personally identifiable information          Browser type          Language preference          Referring site          Date and time of each visitor request          d) Potentially personally-identifiable information (see also Dataset #4)          IP address          e) User Content:          Proposals related content and comments to proposals.</p>
Standards (formats, estimated volume of data)	<p>Formats/Standards:          Storage: database, SQL for CRUD operations          Export of data: CSV, TXT, PDF or Excel File          Interoperability: JSON, PDF          Estimated Volume: Dependent on the number of operations (from Megabytes to Gigabytes)</p>
<b>Data Sharing</b>	
Data Shared	<p>a) Anonymized and clusterized Personal Data:          Age          Neighborhood/Address          Gender          Education          b) Anonymized votes and preferences          c) User Content          Proposals related content          Any other public user generated content</p>
Data sharing purpose (for dissemination or exploitation)	To allow independent research and independent monitoring of the delivery and outcomes of the Pilot of Lisbon in compliance with the Open Data principles established for EMPATIA.
Format of data sharing	<p>Data will be released:          in machine readable format (CSV)          as charts and other info-graphic elaborations (PDF)          images (GIF)</p>

Means of data sharing	On the website <a href="https://www.buergerbudget.wuppertal.de/">https://www.buergerbudget.wuppertal.de/</a> and on the repository hosted on <a href="https://dkan.empatia-project.online">https://dkan.empatia-project.online</a>
Embargo Period	6 months
Open Data license	<a href="https://creativecommons.org/licenses/by-nc-sa/4.0">https://creativecommons.org/licenses/by-nc-sa/4.0</a>
<b>Archiving and preservation</b>	
Data Storage (including backups)	The dataset is preserved in the OneSource datacentre. There are daily backups and weekly backups to a dedicated server.
Preservation periods and where	The collected data in this dataset is kept no longer than one year after the last access to the Service by the Data Collector. The data is kept in the datacenter(s) of OneSource. All personal data regarding this dataset have been destroyed on 31/12/2017.
<b>Data Security</b>	
Physical protection mechanisms	The datacenter where the data is archived and preserved has strict security policies regarding the physical access. The access to the datacenter is restricted to the CTO and CISO and its access is managed through two-level of security: physical key and alarm. All the access is logged in the alarm system and documented in the internal procedures of OneSource. The infrastructure for data archiving is supported by redundant servers and professional storage systems with active mechanisms of redundancy, and protection at the physical level for power supply (UPS systems). Storage uses fiber channel SAN with RAID and multiple servers. The backup system includes automated procedures, to assure data protection at two levels: In-datacenter backups and external-datacenter backups. The in-datacenter backups include daily copies automated and keeping an historic of two years. The external-datacenter backups are performed in a weekly basis and keep an history of two years. All the levels of backup employ encryption technologies and all the access to the backups is controlled in the SIEM of OneSource. Indeed, alerts of level 1 (i.e. marked with high severity) in the SIEM are scaled to the CISO for analysis regarding the possible security threats.
Logical protection mechanisms	The platform implements security mechanisms to protect the data and to manage the access to the data. As the platform is composed by several and independent components, these are distributed in different servers. All the data, especially personal data, is protected through specific components, that only allow access to the data to other components with valid JWT tokens (generated through valid authentication mechanisms). All the tokens have strict policies for expiration (i.e. 10 minutes), requiring new logins to provide valid tokens. As stated, all the transactions that require access to data require valid tokens. As such, no access to data can be performed without a valid login and access permissions, which are set according to the role of the user (e.g. if manager of an entity, or simply as an user participating in the PB process). All the accesses are analyzed in the SIEM of OneSource, where level 1 events (i.e. marked with high severity) are scaled to the CISO, while level 2 (i.e. marked with medium severity) and level 3 (i.e. marked with low severity) are managed by the network administrator and systems operators of OneSource.
<b>Ethical remarks</b>	
Legal framework	International Framework:

	<p>Directive 95/46/EC .  National Framework:  There is a system of regulations enforcing locally the Data Protection. First of all, the website is created in accordance with the provisions of the Telemedia Act (TMG). The processing of personal data is also based on the Telemedia Act. Second, the City of Wuppertal is subject to the Federal State Data Protection Act (DSG NRW). The third party involved in the implementation of the process is subject to the Federal Data Protection Act (BDSG). The Federal State Commission for Data Protection and Information Freedom of North Rhine-Westphalia (Landesbeauftragte für Datenschutz und Informationsfreiheit NRW) is the authority which oversee over controversies regarding the enforcement of Data Protection and other relevant ethical issues. It explicitly shows that the data protection strategies are compliant with local, national and international regulations. The data is passed on to third parties without a user's consent only in case of the legal obligation. In this case, §§ 14-17 DSG NRW shall apply, unless special regulations exist. § 9 DSG NRW applies to the setup of automated retrieval procedures.</p>
Policies /Information Sheet	<a href="https://www.buergerbudget.wuppertal.de/auth/privacyPolicy">https://www.buergerbudget.wuppertal.de/auth/privacyPolicy</a>
Ethical risks	Dataset includes Personal and Sensitive Data. Adequate security measures and protocols shall be implemented by the Data Controller in order to ensure the enforcement of the protection of personal data according to the provisions detailed in the privacy policies.