Big Policy Canvas

D6.1 Data Management Plan

Document Identification			
Status	Final	**Due Date**	31/12/2017
Version	2.0	**Submission Date**	28/12/2017

Related WP	WP6	Document Reference	D6.1
Related Deliverable(s)	N/A	**Dissemination Level (*)**	PU
Lead Participant	NTUA	**Lead Author**	Ariadni Michalitsi - Psarrou
Contributors	Panagiotis Kokkinakos	**Reviewers**	Esther Garrido (ATOS)
			Nuria Rodríguez (ATOS)

Keywords:
Data management plan, metadata, data protection.

Document Information

List of Contributors	
Name	Partner
Ariadni Michalitsi - Psarrou	NTUA
Panagiotis Kokkinakos	NTUA

Document History			
Version	Date	Change editors	Changes
0.1	24/10/2017	NTUA	ToC
0.2	10/11/2017	NTUA	1st draft with inputs from NTUA
0.3	27/11/2017	ATOS	1st draft with inputs from ATOS
0.4	28/11/2017	LISBON	1st draft with inputs from LISBON
0.5	1/12/2017	NTUA	1st draft - Consolidated inputs
0.6	15/12/2017	NTUA, ATOS, FOKUS, LISBON COUNCIL	2nd draft with addressed comments and new inputs
0.7	21/12/2017	NTUA	Version ready for final review
0.8	27/12/2017	ATOS	Reviewer's Comments on the final document
0.9	27/12/2017	NTUA	Final version after reviewer's comments
1.0	28/12/2017	ATOS	FINAL VERSION TO BE SUBMITTED
1.1	10/01/2019	NTUA	Updated references to WP6 Ds
2.0	29/01/2019	ATOS	SECOND FINAL VERSION WITH UPDATES

Table of Contents

List of Figures

List of Acronyms

Abbreviation / acronym	Description
BPC	Big Policy Canvas
CDB	Common Dissemination Booster
DMP	Data Management Plan
DoA	Description of Action
Dx.y	Deliverable number y belonging to WP x
EC	European Commission
EU	European Union
ICT	Information and Communication Technology
PA	Public Administrations
TBD	To Be Decided
WP	Work Package

Executive Summary

The Big Policy Canvas project participates in the Pilot on Open Research Data launched by the European Commission along with the Horizon 2020 programme. Hence, this deliverable sets the project's Data Management Plan, which conforms to the Guidelines on Data Management in Horizon 2020 and specifies the types of data to be generated and collected during the project duration along with the metadata related to them, and the scheme of their archiving and preservation.

The Big Policy Canvas consortium will follow a series of dedicated activities to publish the project outcomes so as to communicate and spread the knowledge to all interested communities and stakeholders and get feedback from them.

The type of data to be generated and collected will be obtained with the collaboration of researchers, the external experts (both the Experts Committee and Experts Advisory Group) and other collaborators. These include:

1. List of Needs
2. List of Trends
3. List of Technological Assets
4. Community Contacts
5. Community Feedback
6. Roadmap
7. Guidelines & Recommendations

In the case where these data will contain personal information - data which relate to an individual who can be identified from those data and/or other information which may come into the possession of any interested stakeholder, and includes any expression of opinion or intention about the individual – Big Policy Canvas will follow the respective, new EU General Data Protection Regulation.

Furthermore, the Big Policy Canvas publication infrastructure consists of several web-based publication platforms that together provide long-term open access to all publishable, generated or collected results in the project: the project's website, ResearchGate, ownCloud, the BPC Knowledge base and other prospective data archiving & publishing infrastructure.

To conclude, this document also addresses the data management process of the project's deliverables. The Big Policy Canvas consortium will follow the same methodology for data sharing, storage and preservation of the forthcoming deliverables, respecting the deliverables' classification as this was defined in the DoA.

1 Introduction

1.1 Purpose of the document

The present deliverable (D6.1) entitled "Data Management Plan" is particularly associated with T6.1 "Dissemination and Communication Strategy" of WP6 and, as such, its main purpose is to document, an initial data management plan for the project, highlighting the project's data archiving and publishing infrastructure and the template under which the project's results will be documented with respect to the management of the data they provide.

Hence, the present deliverable aims to fulfil the following main objective:

- To develop a plan for the data management of the project, identify the infrastructure to be used for data archiving and publishing and list the various expected project's results (from the perspective of the data and information they encapsulate).

1.2 Relation to other project work

WP6 is a horizontal component within the project work plan and aims at supervising the integrity and consistency of all dissemination efforts to achieve the goals mentioned above. In this context, Work Package 6 will retain close collaboration with all project's WPs to ensure that all up to date information and knowledge produced within the project will be effectively recognised and disseminated. Closer connection can be identified though with WP2 "Project Community Establishment, Networking Support and Project's Engagement Activities" that focuses on identifying key stakeholders working in the area of data-driven policy-making and policy-modelling, since WP6 may enforce and facilitate community building, thus these two WPs are closely coupled.
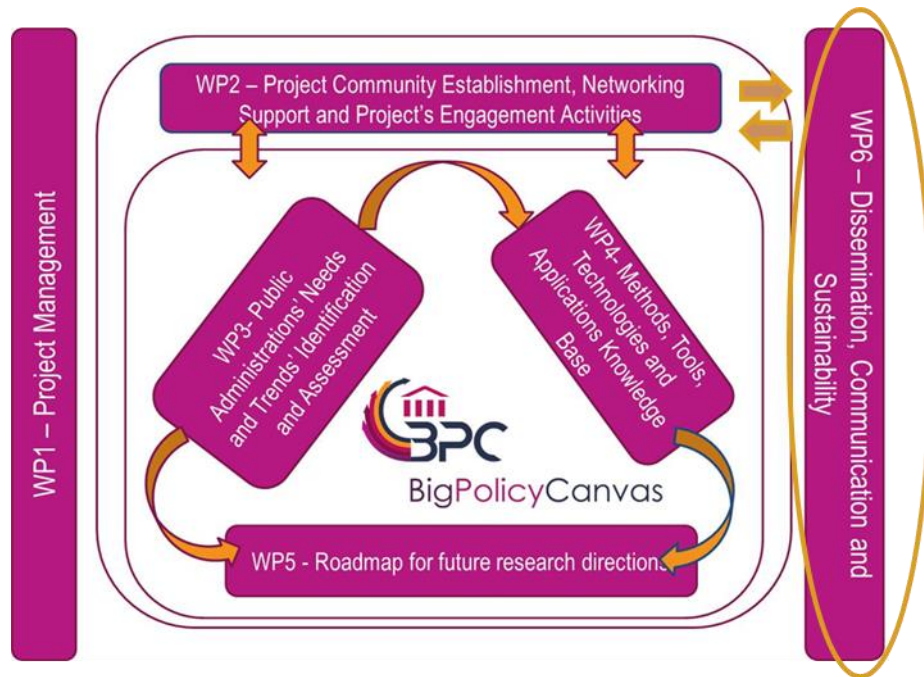
Figure 1-1 Relation of WP6 with the other WPs

1.3 Structure of the document

The rest of this document is structured in the following major chapters:

Chapter 2 refers to the Data Management Plan for the project, exposing the methodological framework that will be used, the data archiving and publishing infrastructure to be exploited and the expected project's results.

Chapter 3 summarizes the main conclusions of the document.

2 Data Management Plan

The Big Policy Canvas consortium will follow a series of dedicated activities to publish the project outcomes so as to communicate and spread the knowledge to all interested communities and stakeholders and get feedback from them. The goal of this section is to define how listing of results and research data that can be published during a research project will be accomplished and describe how these data will be handled from their acquisition and even after the project's end; how these will be collected, processed or generated and following what methodology and standards, whether and how this data will be shared and/or made open, how it will be curated and preserved, etc.

2.1 Methodological Framework for DMP

The general strategy for data management, in accordance with the EC Guidelines for FAIR Data Management in Horizon 2020[1], will be based on the identification and classification of the generated and collected data, the standards and metadata to be used, their exploitation and availability, as well as their sharing, archiving and preservation. In that view, a methodology should be outlined that will make research data generated in the context of the Big Policy Canvas project, findable, accessible, interoperable and reusable.

The Big Policy Canvas DMP aims to cover the whole data life cycle. Hence, the task *T6.1 Dissemination and Communication Strategy and Data Management Plan* in WP6 will be devoted to formulate and continuously evolve the Big Policy Canvas research data management plan in accordance with the H2020 guidelines regarding Open Research Data. In this task, the metadata, procedures and file formats for note-taking, recording, transcribing, storing visual data from participatory techniques, and anonymising semi-structured interview and focus group discussion data will be developed and agreed.

Data Management Plan (DMP) is not required to provide detailed answers to all the questions in this first version. Rather, the DMP is intended to be a living document that will be updated over the course of the project whenever significant changes arise, such as (but not limited to) new data, new innovations, changes in the consortium members and others.

Regarding the **type of data** to be generated and collected, these will be obtained with the collaboration of researchers, the external experts (both the Experts Committee and Experts Advisory Group) and other collaborators.

In the case where these data will contain personal information - data which relate to an individual who can be identified from those data and/or other information which may come into the possession of any interested stakeholder, and includes any expression of opinion or intention about the individual – Big Policy Canvas will follow the respective, new EU General Data Protection Regulation[2].

In what concerns **standards and metadata** to be used, publications (deliverables and papers) will serve as the main piece of metadata for the shared data. Therefore, formats to be used mainly include

---

[1] Guidelines on Data Management in Horizon 2020, European Commission (2016) https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
[2] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:201:0037:0047:en:PDF

.doc, .pdf and .xls files, which substantially reduce the amount of metadata, while other standards do not apply to this project.

In order to decide whether results (i.e. all kind of artefacts collected or generated during the project) should be published or not, a list of questions has been introduced by the project to facilitate their classification as either *public (*under the open access policy) or *non-public*, as follows:

1. *Does a result provide significant value to others or is it necessary to understand a scientific conclusion?*

2. *Does a result include personal information that is not the author's name?*

3. *Does a result allow the identification of individuals even without the name?*

4. *Does a result include business or trade secrets of one or more partners of the project?*

5. *Does a result name technologies that are part of an ongoing, project-related patent application?*

6. *Can a result be abused for a purpose that is undesired by society in general or contradict with societal norms and the project's ethics?*

7. *Does a result break national security interests for any project partner?*

2.2   Data archiving and publishing infrastructure

The Big Policy Canvas publication infrastructure consists of several web-based publication platforms that together provide long-term open access to all publishable, generated or collected results in the project.

In the following subsections, we describe the used platforms.

2.2.1   Project's website

The project's website will be used to provide a short description of the project's objective and its methodology, as well as a short presentation of the consortium. A dedicated page for project's public documents will be available where the most important deliverables of the project will be published in portable document format (.pdf). Furthermore, a blog post section/page will be added to inform the public about events, workshops, news and updates that are relevant to the project's activities. The website will also provide a link to the project's private area, where username/password will be requested from the site in order to upload material and comment on deliverables. The material will be accessed without creating an account, but in order to upload material or comment on deliverables an account will be needed The webpage is hosted by partner Lisbon Council (ipHost provider) at http://www.bigpolicycanvas.eu. All webpage-related data will be backed up on a regular basis. All information on the Big Policy Canvas website will be accessed without creating an account.

2.2.2   ResearchGate

ResearchGate will be used to gather all Big Policy Canvas-related publications and share them with interested researchers in order to further diffuse the research done in the context of Big Policy Canvas.

ResearchGate is expected to raise awareness for project's publications and connect with relevant researchers.

2.2.3 ownCloud

ownCloud will be the project's internal document repository where all the files exchanged within the consortium, including intermediate versions of the deliverables, meetings' material (agenda, notes, presentations, demos, minutes, etc.) and any other document used for gathering inputs from the project's partners will be uploaded. ownCloud is hosted by ATOS and aims to deliver out-of-the-box, collaborative content management, simplifying capturing, sharing, and retrieval of information across virtual teams; boosting productivity; and reducing network bandwidth requirements and email volumes between project team members. Credentials are needed to access any of the ownCloud material, as the platform usage is restricted only to the Big Policy Canvas consortium and to the EC.

Link: https://repository.atosresearch.eu/index.php/apps/files/?dir=%2FBigPolicyCanvas

2.2.4 Knowledge base

The Big Policy Canvas partners will setup a knowledge base that will incorporate all project's findings to produce a repository of value that will facilitate rapid and effective uptake of novel technologies, tools, methodologies and applications that cover the identified by the project public sector needs and exploit available (big) data. This knowledge base intends to act as the project's basic infrastructure that will be constantly updated maintaining the material identified and assessed by Big Policy Canvas, both during and after the end of the project.

2.2.5 Prospective data archiving & publishing infrastructure

Apart from the aforementioned publishing infrastructures, other data archiving and publishing infrastructures, used by other EU-funded research projects and/or suggested by the EC, will be also considered during the following months of the project's duration, in accordance with the resulting needs. Indicatively, examples of such publishing infrastructures that will be examined for their utility in the project are Zenodo, Futurium and JoinUp.

2.3 Project's results

In this section, the datasets used or produced by the Big Policy Canvas partners are listed. In accordance with the EC guidelines for FAIR data management, the necessary information for all the datasets that clarify the way data are collected, documented, stored, preserved and shared are provided.

2.3.1 List of Needs

Dataset Description

Dataset for analysis of existing and emerging public administration's needs is one of the main outcomes of WP3 and in particular T3.1. The data collection techniques to be used will be mainly desk based research, workshops and interviews with public administration representatives and experts from the BPC network. The dataset will be useful for all users of the project outcomes, and will act as well as input for the construction of the roadmap.

Standards and metadata

This dataset is stored in Google Sheets spreadsheets to facilitate the contribution on the identified needs by all consortium members. Information on the need name, need description, need type (e.g. strategical, organisational, technical, etc.) and need source (source name, url and countries on which it is addressed) is being held.

Data Sharing

This dataset will be mainly shared through the WP3, WP4 and WP5 deliverables, which are all public.

Archiving and Preservation (including storage and backup)

The dataset will be preserved in the project internal repository (ownCloud) and in the project website http://www.bigpolicycanvas.eu/. Other archiving and preservation repositories (e.g. Zenodo) will be also examined.

2.3.2 List of Trends

Dataset Description

The dataset regarding the analysis of existing and emerging public administration's trends, along with public administration's needs dataset described above, is one of the main outcomes of WP3 and in particular of T3.1. The data collection techniques to be used will be mainly desk based research, workshops and interviews with public administration representatives and experts from the BPC network. The dataset will be useful for all users of the project outcomes, and will act as well as input for the construction of the roadmap.

Standards and metadata

This dataset is stored in Google Sheets spreadsheets to facilitate the contribution on the identified trends by all consortium members. Information on the trend name, trend description, trend type (e.g. technical innovation, phenomenon, method, concept, etc.), needs addressed by the trend recorded and trend source (source name, url and countries on which it is addressed) is being held.

Data Sharing

This dataset will be mainly shared through the WP3, WP4 and WP5 deliverables, which are all public.

Archiving and Preservation (including storage and backup)

The dataset will be preserved in the project internal repository (ownCloud) and in the project website http://www.bigpolicycanvas.eu/. Other archiving and preservation repositories (e.g. Zenodo) will be also examined.

2.3.3 List of Technological Assets

Dataset Description

The dataset regarding the identification and reporting of methodologies, tools, technologies and applications originating either from public or private sector is being created and reported in the context of WP4 (especially T4.1). Data collection techniques that will be used refer mainly to desk based research, workshops, focus groups with stakeholders met during events and workshops attended by BPC consortium members, interviews with IT experts, from both public and private sector, and interviews and discussions through online communication means with public sector's and policy making experts of the BPC network. The dataset will be useful for all users of the project outcomes, and will act as well as input for the construction of the roadmap.

Standards and metadata

As in the case of the two aforementioned project's results (i.e. list of needs and trends), this dataset is stored in Google Sheets spreadsheets to facilitate the contribution on the identified technological assets by all consortium members. Information on the asset name, asset description, asset type (e.g. tool, database, platform, software, etc.), asset's origin (e.g. public sector, private sector, research domain, etc.), asset application field, needs served by asset recorded and asset source (source name, url and countries on which it is addressed) is being held.

Data Sharing

This dataset will be mainly shared through the WP4 and WP5 deliverables, which are all public.

Archiving and Preservation (including storage and backup)

The dataset will be preserved in the project internal repository (ownCloud) and in the project website http://www.bigpolicycanvas.eu/. Other archiving and preservation repositories (e.g. Zenodo) will be also examined.

2.3.4   Community Contacts

Dataset Description

The community contacts collected during the project's duration will be stored in Excel spreadsheets, the access of which will be restricted to Big Policy Canvas consortium and refer to the following sections and purposes:

- Contact users' register for newsletter subscriptions, containing name and e-mail (both mandatory). This dataset is automatically generated when visitors sign up to the newsletter form available on the project website. The register will be used in order to send issues of the project newsletters.

- Contact user's personal details with regard to messages sent to the website through the Contact form. It includes name, e-mail, message (all mandatory) and (possibly) phone. The contact details will be used to address the inquiry/request and to send information in the scope of the Big Policy Canvas project, after asking for and receiving his/her permission.

- Contacts identified as stakeholders that will build the Big Policy Canvas network and will provide their feedback and support in the dissemination of the relevant to the project information. These stakeholders may be either identified through web sources (e.g. targeted LinkedIn groups) and requested to provide their permission to be considered a Big Policy Canvas network member or may register on their own initiative in the collaboration portal of the project. This dataset will contain their name, surname, job function, domain field/expertise, e-mail contact and location, as well as their project interests or benefits and what they can contribute.

Standards and metadata

This dataset can be imported from, and exported to .doc, .pdf or .xls files

Data Sharing

The mailing list will be used for dissemination and feedback gathering purposes, including disseminating the project newsletter to a targeted audience, inviting community contacts to a Big Policy Canvas event, request their opinion and feedback on a specific topic, etc. An analysis of newsletter subscribers may be performed in order to assess and improve the overall visibility of the project. As it implies personal data, the access to the dataset is restricted to Big Policy Canvas consortium.

Archiving and Preservation (including storage and backup)

The dataset will be preserved in ATOS' servers.

2.3.5 Community Feedback

Dataset Description

Community feedback refers to any kind of feedback produced by the Big Policy Canvas community, either this comes from stakeholders' interviews, questionnaires, focus groups, etc.

Standards and metadata

Regarding interviews and focus groups, data will be collected and stored using digital audio recording whenever interviewees permit it. In any case, the data from these sources will be always held in transcript form in accessible .doc file format (Word). Information coming from questionnaires and any other similar written feedback can be imported from and exported to .doc, .pdf or .xls files.

Data Sharing

These datasets will be used to produce analytical reports on the most important public administration's needs and trends as well as to identify the most appropriate technological assets to tackle these tasks. They will be also used to validate the BPC roadmap and derived guidelines and recommendations. Due to personal data protection, only aggregated information on these datasets will be made accessible, protecting the identity of the engaged stakeholder, if deemed necessary. In case where copyright and IPRs issues are raised, the contributors of feedback will bear the copyright, but they will be asked to assign a Creative Commons License, so that Big Policy Canvas can freely use their contributions, respecting the terms of this license.

Archiving and Preservation (including storage and backup)

These datasets will be preserved in the internal project repository (ownCloud).

2.3.6 Roadmap

Dataset Description

This dataset will be the outcome of the analysis and matching of the identified public administration's needs and trends with the identified technological assets covering specific needs and will provide information on what is already done and what is available at the moment. It is a result coming from WP5 activities taking input from WP3 and WP4.

Standards and metadata

This dataset is a combination of .doc and .pdf documents.

Data Sharing

This dataset will be mainly shared through the WP5 deliverables (D5.1 and D5.2), which are public.

Archiving and Preservation (including storage and backup)

The dataset will be preserved in the project internal repository (ownCloud) and in the project website http://www.bigpolicycanvas.eu/. Other archiving and preservation repositories (e.g. Zenodo) will be also examined.

2.3.7 Guidelines & Recommendations

Dataset Description

Part of WP5 activities is also the elaboration of practical research directions and recommendations to all interested BPC stakeholders. These recommendations and research directions will stem from all the work implemented during the project, especially under WP3, WP4 and WP5 activities, building on the exchange with the community of stakeholders.

Standards and metadata

This dataset is a combination of .doc and .pdf documents.

Data Sharing

This dataset will be mainly shared through D5.3, which is a public deliverable.

Archiving and Preservation (including storage and backup)

The dataset will be preserved in the project internal repository (ownCloud) and in the project website http://www.bigpolicycanvas.eu/. Other archiving and preservation repositories (e.g. Zenodo) will be also examined.

2.4 Data management of other project documents

2.4.1 Project's Deliverables

In this subsection, the data management process of the project's deliverables that have been delivered so far is briefly described. For each of these datasets, the necessary information that characterise the document, describe its content, its format and its metadata, and the way it has been shared and stored, is provided. The Big Policy Canvas consortium will follow the same methodology for data sharing, storage and preservation of the forthcoming deliverables, respecting the deliverables' classification as this was defined in the DoA.

In the following subsections, the so far submitted project's deliverables are listed.

2.4.1.1 D1.1 – Project Management Handbook

Dataset Description

The deliverable defines the structures, the procedures, and the supporting documents that need to be appropriately established in order to assure the quality of the project deliverables and project management activities. It identifies potential risks and a management plan to face these situations.

Standards and metadata

The document is stored in the cross-platform portable document format (.pdf). Metadata will be added manually and include the title, the partner organisations and keywords that classify this report, once the report has been accepted by the EC.

Data Sharing

This document is classified as "Confidential" and thus, access to it is restricted to the consortium and the EC.

Archiving and Preservation (including storage and backup)

The document, as well as all earlier versions of the document, is archived on the project-internal ownCloud repository. The repository is hosted in a server, which is backed on a regular basis by ATOS.

2.4.1.2   D2.1 – Identified Stakeholders & Networking Activities Planning

Dataset Description

This deliverable describes the process of identification and clustering of stakeholders, including the rationale for the involvement of stakeholders in the project and the provisional identification process, as well as the future networking activities of the project, outlining the strategic plan for building the Big Policy Canvas community and the initial version of the community building plan. Furthermore, it contains a preliminary list of communities, related projects and stakeholders.

Standards and metadata

The document is stored in the cross-platform portable document format (.pdf). Metadata will be added manually and include the title, the partner organisations and keywords that classify this report, once the report has been accepted by the EC.

Data Sharing

The document will be published openly on the Big Policy Canvas webpage. The access will be free for everyone and without restrictions.

Archiving and Preservation (including storage and backup)

The document will be published on the Big Policy Canvas webpage. All earlier versions of the document are archived on the project's internal ownCloud repository. The repository is hosted in a server, which is backed on a regular basis by ATOS.

2.4.2   Scientific Publications

With regard to peer-reviewed scientific publications that will result from the project, an Open access publishing approach ('gold' open access) will be followed, according to the Regulation and the Rules of Participation for H2020[3]. All publications will be also become available in ResearchGate to support their further dissemination.

There are no scientific publications to document yet.

---

[3] http://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf

3 Conclusions

The deliverable at hand, entitled "Data Management Plan" is preparatory of the activities to be conducted within WP6 on Dissemination, Communication and Sustainability and details the Data Management Plan for the project, where, apart from the presentation of the methodological framework, the project's outcomes to be produced and disseminated are described with respect to the standards and metadata relating to them, their sharing, their archiving and their preservation. In this context, the main project's data archiving and publishing infrastructure consists of (a) the project's website, where a short description of the project's objective and methodology is presented among other things, (b) OwnCloud, which is the project's internal document repository, (c) ResearchGate, that aims to gather all Big Policy Canvas-related publications and (d) the project's Knowledge Base that will incorporate all project's findings. In what concerns the project's results, as part of tangible data for the project, the list of needs, trends and technological assets are considered, as well as the project's community contacts, their feedback and the project's roadmap and guidelines. Of course, the project's deliverables and scientific publications are also part of the project's results and so they are also being described as such.

References

[1] Big Policy Canvas Description of Action Part B, Coordination and Support Action, Contract Number- 769623, 2017.

[2] Big Policy Canvas deliverable lead author: Francesco Mureddu (LISBON COUNCIL), D2.1 Identified Stakeholders & Networking Activities Planning. Deliverable of the Big Policy Canvas project (2017).