Ref. Ares(2019)2502519 - 09/04/2019

Towards We-Government: Collective and participative approaches for addressing local policy challenges

Grant Agreement number: 693514

Deliverable

D6.4

Data Management Plan

Version 2.3

Project co-funded by the European Commission within H2020-EURO-2014-2015/H2020-EURO-6-2015		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services	
RE	Restricted to a group specified by the consortium (including the Commission Services	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Author List

Organisation	Name	Contact Information
empirica	Lutz Kubitschke	lutz.kubitschke@empirica.com
empirica	Sonja Müller	sonja.mueller@empirica.com

Status, Abstract, Keywords, Statement of originality

Dissemination level:	Public
Contractual date of delivery:	31 January 2019
Actual date of delivery:	09 April 2019
Work Package:	WP6 Overall Project Management
Type:	Report
Approval Status:	Final
Version:	2.3 (revised following final review meeting)

Abstract:

According to available guidance, H2020 projects are to provide an initial Data Management Plan (DMP) within the first six months of the project. The DMP should be updated during the project lifetime. The present document represents an update of the DMP for the WeGovNow project.

The current report starts with a summary of the data collated in the framework of the WeGovNow pilot evaluation. This is followed by a description of how access quantitative data collated for the purposes of formative pilot platform evaluation will be preserved. Next, the resources allocated to this are assessed and measures for ensuring data security are described. This is followed by a description how ethical aspects and data privacy according to GDPR has been addressed. Finally, it is described how access to publications is preserved

Keywords:

Data management plan (DMP), data summary, FAIR data, allocation of resources, data security, ethical aspects.

Statement of originality

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

History

Version	Date	Activity	Revised by
0.1	30/06/2016	Creation of the initial document structure & content	Lutz Kubitschke
0.2	25/07/2016	Main content added	Lutz Kubitschke
0.3	29/07/2016	Peer review	Karsten Gareis
0.4	29/07/2016	Finalisation	Lutz Kubitschke
1.0	29/07/2016	Formatting and submission	Karsten Gareis
1.1	19/07/2016	Document structure revised	Lutz Kubitschke
1.2	24/07/2017	Detailing of the DMP across all sub-sections	Lutz Kubitschke
1.3	26/07/2017	Peer review	Sonja Müller
2.0	28/07/2017	Finalisation	Lutz Kubitschke
2.1	30/01/2019	Previous version updated	Lutz Kubitschke
2.3	09/04/2019	Previous version updated and extended	Lutz Kubitschke

Contents

1 Introduction

According to available guidance, H2020 projects are to provide a first version of the Data Management Plan (DMP) within the first six months of the project[1]. The initial DMP should be updated during the project lifetime. The present document represents a final update of the initial version of the DMP.

The DMP presented throughout this document starts with a summary of the data collated in the framework of the WeGovNow pilot evaluation (Chapter 2). This is followed by a description of how access quantitative data collated for the purposes of formative pilot platform evaluation will be preserved (Chapter 3). Next, the resources allocated to this are assessed (Chapter 4) and measures for ensuring data security are described (Chapter 5). This is followed by a description how ethical aspects and data privacy according to GDPR has been addressed (Chapter 6). Finally, it is described how access to publications is preserved (Chapter 7).

2 Data summary

2.1 Administrative information

Grant Agreement No.: 693514

Acronym: WeGovNow

DPM version: 2.3

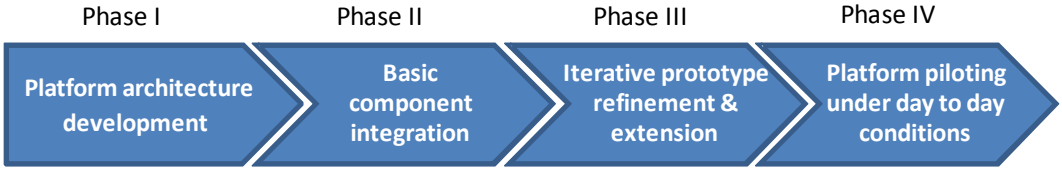
Planned update: n. a.

DMP responsible: Lutz Kubitschke

2.2 Purpose of data collation

The WeGovNow project is to develop a new type of online platform for citizen engagement, thereby integrating existing software applications and newly developed ones in accordance with a multi-staged development approach (Figure 1).

Figure 1 – Multi-staged development approach of the WeGovNow online platform



[1] European Commission, Directorate-General for Research & Innovation: Guidelines on Guidelines on FAIR Data Management in Horizon 2020, Version 3.0, 26 July 2017

The WeGovNow platform will has been piloted in three municipalities from project month 25 onwards. During the pilot evaluation phase data have been gathered in relation to different evaluation objectives, namely whether

a) the WeGovNow pilot system works as anticipated (viability perspective)
b) the WeGovNow pilot system is worth being maintained after the ending of the pilot duration (sustainability perspective)

Outcomes have been feed into the development of guidance on the further mainstreaming of the WeGovNow pilot platform after the ending of the pilot project duration, including recommendations on further exploitation of the WeGovNow pilot platform as a major project output.

2.3 Type and format of data collated

Different types of evaluation data have been gathered throughout the project's pilot phase, including quantitative data:

- Quantitative evaluation data generated by means of automated monitoring of platform usage:
  The WeGovNow online platform developed includes different functional software components which enable the pilot users to post and manipulate different types of content. This is reflected by the case-based data set including the following variables:
  - User account created
  - User account automatically validated
  - User account manual validation requested
  - User account manually validated
  - Date of birth of registered user account holder
  - Sex of registered user account holder
  - No. of original contributions posted through the WeGovNow First Life platform component
  - No. of comments posted through the WeGovNow First Life platform component
  - No. of original contributions posted through the WeGovNow Community Maps component
  - No. of original contributions updated through the WeGovNow Community Maps component
  - No. of comments posted through the WeGovNow Community Maps component
  - No. of objects created through the WeGovNow First Life component
  - No. of objects updated through the WeGovNow First Life component
  - No. of objects deleted through the WeGovNow First Life component
  - No. of votes casted d through the WeGovNow LiquidFeedback component

  These data have been processed in a common spread sheet processing programme format (Microsoft Excel).

- Qualitative evaluation data gathered by means of different data gathering methods: Semi-structured interviews were conducted with a number of selected stakeholders at the pilot sites for the purposes of the viability assessment, focussing on identifying potential impacts of the pilot service on the various stakeholder groups involved. Outcomes were documented in terms of a common reporting template. Further to this qualitative evaluation data has been gathered by means focus group sessions. In thematic regard, these events focused on gathering the stakeholders' perceptions of the utility, usability and reliability of the pilot services, including the technical infrastructure through which the pilot services are delivered at each of the pilot sites. Also, perceived impacts have been addressed. Outcomes were again documented in terms of a common reporting template. All reporting templates were processed as a text files with help of common word processing software (MS Word).

2.4 Re-use of existing data

The WeGovNow project specifically has developed and piloted a new type of civic engagement platform integrating various participation functions. There was no existing data available that could be utilised for the purpose of pilot platform evaluation in the framework of WeGovNow.

2.5 Expected size of data

The following data volume was generated:

- Quantitative data gathered:
  A case-based variable set was generated for 11.833 instances overall.
- Quantitative data gathered:
  Overall, 18 aggregated reporting templates have been processed.

2.6 Expected data utility

Data gathered by means of the project's multi-method evaluation approach specifically relate to the utilisation of the newly developed WeGovNow pilot platform in three participating pilot municipalities. These data are of utility primarily for formative evaluation purposes to support the further mainstreaming of the pilot platform, the latter representing a key output of the project. The pilot platform may however undergo further optimisation prior to envisaged mainstreaming after the ending of the pilot project. Due to the formative nature of the evaluation design adopted for the purpose of the WeGovNow project, the long term value of the evaluation data to be generated can be assessed as very low. Type and nature of the evaluation data gathered do not offer opportunities for subsequent research, e.g. in terms of secondary analyses or replication of results. [2] However, the quantitative data

[2] As pointed out in the literature, in many fields of study there are examples of scientific investigations that cannot be replicated in terms of using independent investigators, methods, data, equipment, and protocols.

set that has become available from automated usage monitoring during the overall project's pilot phase may potentially be utilised by others for verifying any published results referring to these data, albeit with a rather low probability given the strong RTD-focus of the current project.

3 Preserving access to quantitative WeGovNow platform evaluation data

3.1 Making quantitative WeGovNow platform evaluation data findable

Beyond the project duration, it is planned to make an anonymous case-level data set from the automated pilot platform usage monitoring available through the GESIS Datorium repository. The data set consists of 11.833 instances and 11 coded variables. To this end, the data set is intended to be made available after the formal closing of the project.

The user ID which might in conjunction with registration data stored in the pilot platform, at least theoretically, enable the identification of an individual who's platform usage activities have been monitored is not included in the quantitative data set that has been generated for formative evaluation purposes. Also, no person related information stemming from the required registration of the pilot users to the pilot platform has been derived from the platform for evaluation purposes. Hence, no personal data or any other information that would theoretically allow a third party to infer a pilot user's identify has been included in the data set generated for formative evaluation purposes. This ensures anonymity of the data set to be publicly preserved.

Findability will be supported by relying on the metadata framework applied by the data sharing repository to be utilized (GESIS Datorium)[3]. In particular, the following metadata will be provided:

1. Title
2. Principle investigator & Institution
3. Publisher
4. Publication Year
5. Availability
6. Subject Area
7. Topic Classification

In contrast to "replicablility" in this sense, we understand "reproducibility" as making available data sets to others for verifying published results. C.f. e.g. Peng, R. (2009): Reproducible research and Biostatistics. In: Biostatistics, 10 (3): 405-408.

[3] This meta data framework is compatible with the codebook standard of the Data Documentation Initiative (DDI 2). c.f. Wolfgang Zenk-Möltgen and Monika Linne (2014): Metadatenschema zu datorium - Data Sharing Repositorium. GESIS-Technical Reports 2014/103

8. Abstract
9. Geographical Area
10. Data collection Mode
11. Survey Period
12. Rights
13. Notes
14. File description
15. Research Data Type
16. Language
17. Number of Variables
18. Number of Units
19. Unit Type
20. Software

These metadata will be fed by means of a persistent digital identifier (DOI) into the international community of DataCite and thus become findable at an international level.[4]

3.2 Making quantitative WeGovNow platform evaluation data openly accessible

As mentioned above, an anonymous case-level data set derived by means of automated pilot platform usage monitoring will be uploaded to the GESIS Datorium data sharing repository. An account has already been registered.

The data set will be made available in terms of an OpenDocument Spreadsheet (.ods). In conjunction with the metadata to be provided the data will be immediately usable with help of commonly available spreadsheet software. No particular documentation of access software will thus be required.

Access to this data is planned to be provided to everybody. Setting up a Data Access Comitee, e.g. for concluding particular data access agreements, is considered not necessary.

3.3 Making quantitative WeGovNow platform data interoperable

As discussed earlier the type and nature of the data set generated for the purposes of for formative evaluation of the WeGovNow pilot platform is not expected to be of high long-term value (c.f. 2.6). When it comes to potentially re-combinations with different datasets from different origins in particular, no value is seen at all due to the formative evaluation approach pursued.

As discussed earlier as well, the quantitative data set that has become available from automated usage monitoring might be utilised by others for verifying any published results referring to these data, albeit with a rather low probability given the strong RTD-focus of the current project. This will be supported by a documentation of the all variables and numeric

[4] c.f. https://www.gesis.org/fileadmin/upload/institut/presse/Infos___Downloads/Lep_datorium_final.pdf (latest access: 05/04/20

codes included in the data set that will be provided on the data sharing repository (GESIS Datorium). Where possible available standard codes have been followed, for example in the case of identification of date of birth data were coded according to ISO 8601.

3.4 Increase re-use of quantitative WeGovNow platform evaluation data

The data set includes automatically derived monitoring data on pilot platform utilization. Prior to further processing, the data derived from the pilot platform underwent quality assurance procedure in relation to data integrity in terms of a plausibility assessment.

The data set will be made publicly available on GESIS Datorium following the closing of the WeGovNow project. No embargo period is foreseen.

No particular restrictions on the re-use of the data are planned to be imposed. The appropriateness of different licensing agreements under the Creative Commons licensing framework is currently explored.

The data set will remain re-usable until the repository withdraws or goes out of business.

4 Allocation of resources

Costs for data preparation and documentation are covered by the project budget. The costs for data preparation to be FAIR cannot be exactly specified at the current stage. However, expenses for data set preparation, data management and additional documentation concerning those data to be made openly accessible are estimated to not exceed 0.5 person months. No additional expenses are expected to accrue for purchasing supportive tools, e.g. for working with DDI, and repository charges for data submission.

The lead partner (empirica) of the project's evaluation work package (WP4) takes responsibility for data management. Lutz Kubitschke and Sonja Müller are responsible for data storage, archiving and publication.

5 Data security

During the project, all evaluation data has been stored at the server of the lead organization (empirica) responsible for the evaluation work package (WP4), with daily backup at an institutional off-site server. The team member responsible for storage is supported by empirica's IT team. Back up are checked manually at two weeks intervals. No additional costs are accruing for storage and back-up.

When it comes to quantitative monitoring data, sensitive data (the user account ID recorded by the pilot platform) has been separated to create an anonymised data set. Beyond the user ID , no data item processed and stored is assessed as sensitive.

When it comes to qualitative evaluation data generated for evaluation purposes, this data is planned to be stored locally at empirica's servers for 10 years, whereby no costs will be associated with local storage. As described above, qualitative raw data will not be made

available to external parties preventing data privacy threats. Only to project members will be granted access on request, with clearance of a non-disclosure agreement.

6 Ethical aspects

Informed consent for the sharing of anonymised evaluation data and long term preservation was included during data collation. Sensitive data was separated and kept secure.

When it comes to the processing of personal data in the framework of the local validation trials which have been implemented in the three WeGovNow pilot municipalities GDPR compliance was ensured. During the overall project's pilot phase, the WeGovNow platform has been operated by three municipalities under day-to-day, namely in Turin, in the London Borough of Southward and in San Doná di Piave. With help of the platform, each pilot municipality offered a publicly available pilot service to its citizens until the end of the formal project duration. In technological respect, the individual platform components (software applications) were hosted by different consortium members according to a Software-as-a-Service (SaaS) deployment model. Pilot users entered data into the pilot platform through a common web interface. User data were shared across the platform components, either directly between individual components or by means of a common logger data base.

From a legal point of view, two different perspectives deserved attention in framework of the WeGovNow pilots when it comes to GDPR requirements, namely an external one and an internal one:

- External: One the one hand, a legal relationship was established between the pilot service provider (pilot municipality) and the pilot service user (citizen).
- Internal: On the other hand, a legal relationship was established between the pilot service provider (pilot municipality) and those parties processing data on its behalf in accordance with a SaaS deployment model (technology partners hosting one or more platform component remotely).

As far as the processing of personal data was concerned, the GDPR creates obligations for the WeGovNow municipalities offering the pilot service ("data controllers") to their citizens ("data subjects") with help of WeGovNow component providers ("data processors"). These obligations were mat in various ways:

- Terms of use statement (ToU) complying with GDPR requirements was developed by each pilot site and made available through the local pilot platform instances to the pilot users. Consent was requested prior to user registration.
- A data privacy statement was developed by each pilot site and made available through the local platform instances to the pilot users. Consent was requested prior to user registration. In this context, user were informed about which types of personally identifiable information was collected about them across the WeGovNow platform components, how the data is used and how users can control the information that is gathered. Current data protection legislation as well as the new

GDPR put an obligation on data controllers to ensure data subjects can rectify remove or block incorrect data about themselves. Users were also informed about their various rights in relation to data protection as stipulated by GDPR.

- To be able to respond to user request in relation to these rights, it was identified in advance what personal data were held within the individual WeGovNow platform components, where it comes from, who it is shared it with, how its processing can be restricted and how it can be erased. Also, a process was identified din advance how the consortium would react if a pilot user asked to have his/her personal data deleted, for example. In such a case the pilot municipality receiving such a request from one of its citizens was able to rely upon a commonly agreed procedure for informing the partners concerned, monitor how the user claim is met and provide informed feedback on this matter to the pilot user.
- The GDPR makes privacy by design an explicit legal requirement under the term 'data protection by design and by default'. As a general rule, personal data was processed within and across the WeGovNow platform software components only for those purposes intended to be achieved by the component.

When it comes to the legal relationship between the pilot municipalities and the technical partners hosting one or more software components remotely, GDPR differentiates between the "data controller" (the pilot municipalities) and the "data processor" (the technical partners). Processing of personal data on behalf of a data controller requires an assignment in writing between both parties according to GDPR. Therefore a data processing agreement was concluded bilaterally between each of the three pilot municipalities and each WeGovNow component provider prior to the starting of the local pilots. Overall, 12 data processing agreements were hence concluded.

7 Preserving access to publications

Project partners have relied on research results from WeGovNow for authoring scientific papers for journals and book chapters, as well as presenting conference papers in relevant disciplines. To preserve their accessibility these have been made available on different repositories as follows:

- The Zenodo open access repository
- The open access repository of the University of Heidelberg
- The open access repository of the University of Turin
- The Liquid Democracy Journal (permanently archived at German National Library)
- The ACM digital Library
- IEEE Xplore Digital Library
- International Conference on Cartography and GIS (ICC&GIS)

Title (Year)	Authors	Journal/book/conference	Type	Link to publication

Title (Year)	Authors	Journal/book/conference	Type	Link to publication
A Fair Distance Function (2017)	Behrens, J. and B. Swierczek	The Liquid Democracy Journal on electronic participation, collective moderation, and voting systems	Other	http://www.liquid-democracy-journal.org/issue/5/
LiquidFeeback's Issue Limiter (2017)	Behrens, J., Nitsche, A. and B. Swierczek	The Liquid Democracy Journal on electronic participation, collective moderation, and voting systems	Other	http://www.liquid-democracy-journal.org/issue/5/
Unified User Management with LiquidFeedback (2018)	Behrens, J. and B. Swierczek	The Liquid Democracy Journal on electronic participation, collective moderation, and voting systems	Other	http://www.liquid-democracy-journal.org/issue/6/
Data Quality Concept for e-Government Web-Map Based Services (2018)	Noskov A., Zipf A. and A. Rousell	Proceedings 7th International Conference on Cartography and GIS	Conference proceedings	https://iccgis2018.cartography-gis.com/7ICCGIS_Proceedings/7_ICCGIS_2018%20(34).pdf
Computer Vision Approaches for Big Geo-Spatial Data: Quality Assessment of Raster Tiled Web Maps for Smart City Solutions (2018)	Noskov A.	Proceedings 7th International Conference on Cartography and GIS	Conference proceedings	https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/noskov2018rastertilesqualityinitial.pdf
Open Source Tools for Coastal Dynamics Monitoring (2018)	Noskov A.	Proc. SPIE 10773, Sixth International Conference on Remote Sensing and Geoinformation of the Environment	Conference proceedings	https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/noskov2018osrccoastdyn.pdf
Backend and Frontend Strategies for Deployment of WebGIS Services (2018)	Noskov A. and A. Zipf	Proc. SPIE 10773, Sixth International Conference on Remote Sensing and Geoinformation of the Environment	Conference proceedings	https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/noskov2018fbswebgis.pdf
Smart City WebGIS Applications: Proof of Work Concept for High-level Quality-of-Service Assurance (2018)	Noskov, A.	ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-4/W7	Conference proceedings	https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/noskovzipf2018pow.pd

Title (Year)	Authors	Journal/book/conference	Type	Link to publication
				f
Definition of Contour Lines Interpolation Optimal Methodas for E-Government Solutions (2018)	Noskov, A. and A. Zipf	ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-4/W8	Conference proceedings	https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/noskovzipf2018interp.pdf
Modelling and Assessing Spatial Big Data: Use Cases of the OpenStreetMap Full-History Dump (2019)	Noskov, A. et al.	Spatial Planning in the Big Data Revolution	Book chapter	http://ar.n-kov.com/i/NoskovGrinbergerPapapesiosRousellTroiloZipf2019LowLevelFHD.pdf
Open-Data Driven Embeddable Quality Management Services for Map-Based Web Applications (2019)	Noskov, A. and A. Zipf	Big Earth Data	Journal Article	https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/noskovzipf2019embquality.pdf
From E-Government to We-Government: an analysis towards participatory public services in the context of the H2020 WeGovNow (2018)	Tsampoulatidis, I., Kompatsiaris, I. and N. Komninos	Information Society and Smart Cities Conference University of Cambridge, United Kingdom	Conference proceedings	https://zenodo.org/record/2578929#.XJoSoKAiFLw
La Pubblica Amministrazione responsabile : un caso di digital welfare (2018)	Visentin, M. and G. Antonini	Rivista Italiana di Public Management	Journal Article	https://zenodo.org/record/2579141#.XJoSXaAiFLx
First Life, the Neighborhood Social Network: a Collaborative Environment for Citizens (2016)	Antonini, A. et al.	Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion	Conference publication	https://iris.unito.it/handle/2318/1646139#.XHQHFaAiFhE
WeGovNow: a map based platform to engage the local civic society (2018)	Boella, G. et al.	WWW '18 Companion Proceedings of the The Web Conference 2018	Conference publication	https://dl.acm.org/citation.cfm?id=3191560
WeGovNow: an integrated platform for social engagement in shaping future cities (2018)	Boella, G. et al.	4th Italian Conference on ICT for Smart Cities And Communities 2018	Conference publication	https://iris.unito.it/handle/2318/1693782#.XHP_aqAiFhE
Back to public: Rethinking	Lupi, L. et al.	Proceedings of the	Conference	https://iris.unito.it/handle/231

Title (Year)	Authors	Journal/book/conference	Type	Link to publication
the public dimension of institutional and private initiatives on an urban data platform (2016)		2016 IEEE International Smart Cities Conference (ISC2)	publication	8/1646137#.XHQHiKAiFhE
MiraMap: A We-Government Tool for Smart Peripheries in Smart Cities (2016)	De Filippi, F. at al.	IEEE Access Journal	Journal Article	https://ieeexplore.ieee.org/document/7444140