ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

PERCEIVE project - Deliverable D8.5 "PERCEIVE (Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe): Data Management Plan. Version (M6)"

This is the submitted version (pre peer-review, preprint) of the following publication:

Published Version:
Mollona, E., Pareschi, L., Reverberi, P., Brasili, C. (2017). PERCEIVE project - Deliverable D8.5 "PERCEIVE (Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe): Data Management Plan. Version (M6)" [10.6092/unibo/amsacta/5551].

Availability:
This version is available at: https://hdl.handle.net/11585/605010 since: 2020-02-24

Published:
DOI: http://doi.org/10.6092/unibo/amsacta/5551

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

12 October 2024

PERCEIVE

Perception and Evaluation of Regional and Cohesion Policies by Europeans and Identification with the Values of Europe

Call	**H2020-REFLECTIVE-SOCIETY-2015**
Topic	**REFLECTIVE-3-2015: European cohesion, regional and urban policies and the perceptions of Europe**
Grant Agreement number	**693529**
Work package	**WP8 - Project Management**
Task	**8.4 - Managing the knowledge generated by the project and the IPRs**
Deliverable	**8.5 - Data Management Plan**
Lead beneficiary	**ALMA MATER STUDIORUM - Università di Bologna (UNIBO)** **contact: Prof. EDOARDO MOLLONA, edoardo.mollona@unibo.it**
Authors	**UNIBO** **ALL PARTNERS HAVE CONTRIBUTED TO THIS DELIVERABLE**
Version	**0.4 Final**
Approval Status	**approved by Coordinator and Partner Team Leaders**
Planned delivery date	**28/02/2017 (M6)**
Actual delivery date	**28/02/2017 (M6)**
Dissemination level	**Confidential (only for members of the consortium)**
Project website	**http://www.perceiveproject.eu/**
Duration	**01 September 2016 - 31 August 2019 (36 months)**

The information in this document reflects only the author's views. The Research Executive Agency of the European Commission is not responsible for any use that may be made of the information it contains. The user therefore uses the information at its sole risk and liability.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement number 693529

Document History

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
0.1	03/02/2017	UNIBO	First draft
0.2	08/02/2017	UNIBO	Revised draft by Coordinator team
0.3	27/02/2017	UNIBO	Revised draft by Partners
0.4	27/02/2017	UNIBO	Final version

Scheduled Data Management Plan (DMP) updates

The DMP is a document which evolves during the lifespan of the project and registers all relevant changes in the life-cycle of all the research data sets of PERCEIVE project. Updated versions of PERCEIVE DMP are already planned (see table below). Moreover, this document will be updated whenever important changes in the data or the data management policy occur.

Version	Expected by project month (M)
RV1	14
RV2	36

Partner Acronyms

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA (Italy)	**UNIBO**
GOETEBORGS UNIVERSITET (Sweden)	**UGOT**
INSTITUTE OF AGRICULTURAL ECONOMICS (Romania)	**IEA**
INSTYTUT EKONOMIKI ROLNICTWA I GOSPODARKI ZYWNOSCIOWEJ-PANSTWOWY INSTYTUT BADAWCZY (Poland)	**IAFE-NRI**
UNIVERSITAT DE BARCELONA (Spain)	**UB**
UNIVERSITY OF PORTSMOUTH HIGHER EDUCATION CORPORATION (United Kingdom)	**PBS**
WIRTSCHAFTSUNIVERSITÄT WIEN (Austria)	**WU**
SOCIETA'COOPERATIVA BAM! STRATEGIECULTURALI (Italy)	**BAM!**

Table of Contents

The Data Management Plan (DMP) 1
1. Data summary 1
2. FAIR Data 3
2.1 Making data findable, including provisions for metadata 3
2.2 Making data openly accessible 4
2.3 Making data interoperable 7
2.4 Increase data re-use (licensing) 7
2.5 Allocation of resources 8
2.6 Data security 10
2.7 Ethical aspects 11
3. Data sets overview 13
Annex I: Data sets tables 14
WP1 - Framework for comparative analysis: differences, implementation, perceptions and data gathering 14
WP2 - Evaluation of EU citizens' perception of the EU project in relation to regional performance of the Cohesion Policy and institutional quality 17
WP3 - The effectiveness of projects' communication strategies 20
WP4 - Spatial determinants of policy performance and synergies 22
WP5 - The multiplicity of shared meanings of EU and Cohesion Regional and Urban Policy at different discursive levels 25
WP6 - Virtual learning environment 27

The Data Management Plan (DMP)

This DMP provides details regarding all the research data collected and generated within PERCEIVE project. In particular, it explains the way research data are handled, organized, licensed and made openly available to the public, and how they will be preserved after the project is completed. This DMP also provides motivations when versions or parts of the project research data cannot be openly shared on account of third-party copyright issues, confidentiality or personal data protection requirements, or when open dissemination could jeopardize the project achievements.

This DMP reflects the current state of the art of PERCEIVE project. The details and the final number of the project data sets may vary during the course of research. The variations will be recorded in updated versions of this DMP.

1. Data summary

PERCEIVE aims at understanding what factors influence the perception of Europe at a regional level and the reasons why there are variations in Cohesion Policy's results, in citizens' appreciation and identification with the European Union.

To achieve this objective, PERCEIVE will analyze the relationship among European Cohesion, Regional and Urban Policies and the Perceptions of Europe.

In particular, PERCEIVE will develop a comprehensive theory of "cohesion in diversity" and use this theory to create a better understanding of the channels through which European policies contribute to create:

- different local understandings of the EU;
- different levels of European identification across deeply different European regions.

PERCEIVE will use qualitative and quantitative analytical methods, such as surveys, focus groups, case studies and econometric modelling. In addition, it will also use particularly innovative methods, such as quantitative discourse analysis, to elicit meaning structures in public discourse about the EU, its regional policy and European identity.

Moreover, PERCEIVE will produce a computer simulation environment and will embed it into a virtual platform that cohesion policy stakeholders will be able to use to produce "what-if" analysis and long-term scenario analysis of the effects of policies.

The project will generate and collect several different types of research data: mainly numerical and textual data, but also audio and modeling data. Research teams will convert data from proprietary formats and make them available in well-known and documented open formats to allow accessibility, reusability and long term preservation (see Table 1 for details).

Table 1
Summary of data format

Type of data	Formats used during data processing	Formats for sharing reuse and preservation
Numerical or textual tabular data	Microsoft Excel (.xls/.xlsx)	Comma-separated values (.csv)
Qualitative textual data	Microsoft Word (.doc/.docx)	Rich Text Format (.rtf) or text (.txt)
Audio data	mp3 format (.mp3)	*Audio recordings will be deleted after their transcription and only the processed transcripts will be shared and preserved.*
Topic modeling data	Mallet format (.mallet)	Comma-separated values (.csv)
Simulation model data	Text model format (.mdl)	*The mathematical model will be saved using standard differential equations symbols in .rtf, .csv and .txt files because the simulation model is developed using a proprietary software. Simulated values will be saved as numerical data, as specified above.*
Statistical data	STATA format (.dta)	Comma-separated values (.csv), Stata format (.dta)

Documentation files explaining all relevant details regarding data collection, processing methodologies and quality assurance will be deposited in institutional or public repositories along with the data sets in .odt, .rtf or .pdf formats.

The project will reuse a variety of existing data from different sources: data on the Cohesion Policy projects funded by Structural Funds in the programming period 2007-2013 managed by Local Managing Authorities; data from EU and government databases (such as Eurobarometer[1], Eurostat[2], Urban Audit data set [3] and databases from the ONS, the Office for National Statistics[4]); data from communication sources of different genres produced by the EC (e.g. policy documents, explanatory brochures and reports, magazines); descriptions and reports created by Local Managing Authorities; communication materials from newspaper articles, tweets, and blog posts.

The expected size of the data is still uncertain at this early stage of the project but it is estimated that it may not exceed 5 GB.

The data produced can be of interest to different potential users. They include: researchers, students, policy makers, stakeholders, practitioners working on assessment of Cohesion Policy or interested in understanding how the concepts of Cohesion Policy, Europe and European identity are shaped in the public sphere and how to communicate policies effectively and efficiently. The data could also be used as a source for topic-related studies, comparisons and for different analyses, for example for simulating and exploring the scenario that follows alternative specific communication policies.

[1] Eurobarometer, http://ec.europa.eu/public_opinion/index_en.htm
[2] Eurostat, http://ec.europa.eu/eurostat
[3] Urban Audit data set (part of Eurostat), http://ec.europa.eu/eurostat/web/cities
[4] Towns and Cities Analysis database of ONS, https://www.ons.gov.uk/releases/townsandcitiesanalysis

2. FAIR Data

This DMP follows the EU guidelines[5] and describes the data management procedures according to the FAIR principles[6]. The acronym FAIR identifies the main features that the project research data must have in order to be findable, accessible, interoperable and re-useable, allowing thus for maximum knowledge circulation and return of investment.

2.1 Making data findable, including provisions for metadata

At the moment of publication of project results, each research teams will deposit and describe the relative underlying data set(s) in institutional or public data repositories that can attribute persistent unique identifiers to the deposited items. In particular, the UB repository attributes a Handle as persistent identifier to the deposited data sets. The other repositories can mint valid DOIs (Digital Object Identifiers) to identify the deposited data sets. Partners are strongly recommended to use the persistent unique identifiers (DOI or Handle) to cite the data sets as underlying data within their research publications.

The chosen data repositories support standard descriptive metadata to ensure data sets indexing and discoverability. In particular, they all support Dublin Core[7] and DataCite Metadata Schema[8]. Moreover they comply with the OpenAIRE 3.0 requirements for data archives[9]. As a consequence the project data sets will be visible through the OpenAIRE portal[10], facilitating project reporting procedures.

(See Table 2 for the list of the chosen data repositories).

Specific keywords derived, when possible, from thesauri and controlled vocabularies will be associated to each data set to enhance semantic discoverability. The categories described by the European Commission in Annex II of Commission Regulation (EC) No 1828/2006[11] will be used as controlled vocabulary to describe the variables of the EU funded projects analyzed by PERCEIVE.

All relevant documentation explaining data collection procedures and analysis (such as codebooks, methodologies, etc.) will be made available along with the data, in order to guarantee intelligibility, reproducibility and the validation of the project findings. All data sets will be described using standard metadata - such as Dublin Core[7] and DataCite Metadata Schema[8] - and according to the OpenAIRE guidelines[9] in order to ensure metadata interoperability for data sets indexing and discoverability.

[5] Guidelines on FAIR Data Management in Horizon 2020 (Version 3.0, 26 July 2016), http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
[6] The FAIR data principles (Force11 discussion forum), https://www.force11.org/group/fairgroup/fairprinciples
[7] Dublin Core Metadata Initiative, http://dublincore.org/
[8] DataCite Metadata Schema 3.1, https://schema.datacite.org/meta/kernel-3/
[9] OpenAIRE Guidelines for Data Archives, https://guidelines.openaire.eu/en/latest/data/index.html
[10] OpenAIRE, https://www.openaire.eu/
[11] Commission Regulation (EC) No 1828/2006 (pg.47), permalink http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02006R1828-20111201

PERCEIVE research data are organized in data sets, which are named collections of data units with the same focus and scope. This DMP identifies the following common rules for **data set naming** in order to improve data visibility, discoverability, citation and permanent online tracking. The recommended title for each data set consists of:

> *PROJECT ACRONYM: WPnumber: WP title or description specifying WP aims: Tasknumber: Task title or description specifying Task aims: additional information specifying coverage and nature of data (if necessary): version number (in case of revisions or updates)*

Example:

> *PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.1: The framework for the comparative analysis: v.02*

The version number of the data set will be added at the end of the title in case of data revisions to help identifying the data set updates especially in repositories that do not track versioning automatically (see *Annex I* for data set names, unique identifiers and descriptions).

The DMP recommends also the following rules for **file naming**:

- for data set file(s)

  DATASET_PERCEIVE_WPnumber_Tnumber_coverage or other content specifications_date (YYYYMMDD)_vn.file extention

- for the data set relevant documentation explaining data collection procedures and analysis (such as codebooks, users' manuals, methodologies, etc.)

  README_PERCEIVE_WPnumber_Tnumber_coverage or other content specifications_date (YYYYMMDD)_vn.file extention

WPnumber means "work package number" *Tnumber* is "task number", and *vn* is the "version number" (in case of data revisions or updates).

2.2 Making data openly accessible

As a guiding principle, PERCEIVE seeks to make research data openly available, whenever possible, in order to allow dissemination, validation and re-use of research results. To this purpose, all the files will be converted to standard and well-documented open formats and the data sets will be deposited together with all relevant documentation and explanation.

Restrictions on data access or impossibility to share them will be considered only in the following cases:

- when collected data belongs to third party which have denied permission for sharing them on account of confidentiality and proprietary issues;
- protection of personal data of key informants involved in surveys, focus groups, interviews, and case studies;
- when availability of the data would mean that the project's main aim might not be achieved (reasons will be explained in the accessibility details relating to each data set described in *Annex I*);

- other legitimate reasons (that will be explained in the accessibility details relating to each data set described in *Annex I*).

As a consequence, all possible and legitimate actions and strategies will be adopted to allow data sharing including:

- obtaining copyright permissions from third party data owners to be allowed to re-use, reproduce and distribute the collected data;
- converting the files to standard open formats;
- providing all relevant documentation and explanation for the data and the data sets;
- obtaining the consent of stakeholders involved in focus groups and anonymizing and aggregating the data of interviews;
- in case of copyright on raw data derived, collected or elaborated from pre-existing databases or from other original sources (i.e. papers, journal articles, book chapters, reports, video and audio sources), collected data will be made available if the reproduction and sharing are allowed by expressed permission of the right holders or by applicable copyright exceptions and exemptions. Specifically, reproductions and communication of brief excerpts of texts and of other protected works are permitted for illustration purposes for scientific research, provided that the source, including the author's name, is acknowledged and provided that the use does not conflict with the exploitation of the original source and does not unreasonably prejudice the legitimate interests of right holders. Otherwise, only aggregate data resulting from the analysis will be openly published. Anyway, when the sources are freely available on-line in their original repositories, but direct reproduction is not allowed, a detailed account on how the data set was created from the original data will be provided, together with the specification of open repositories from where the original data sets are available. Raw data consisting in full-texts will not be made available without copyright holders permission.

For data that fall under some of the restrictions described above and for which it is not possible to take any action to make them shareable, EU allows complete closure or restricted access to them. PERCEIVE DMP indicates the versions or parts of the data sets that cannot be freely shared providing the specific motivations in *Annex I*.

At the time of publication of results, researchers will deposit the project data that can be shared in a data repository, in order to guarantee their discoverability, access and preservation beyond the project end.

The data repositories chosen by PERCEIVE partners are both institutional and public repositories. They guarantee long term preservation and attribute persistent unique identifiers to the archived data sets (such as DOI or Handle). They support open licenses and different access levels. Finally, they adopt descriptive metadata standards as required by the OpenAIRE Guidelines and allow cross-linking between publications and the relevant data sets.

Each different data set is deposited by the team that is responsible for the data collection and management in the repository of their choice.

Table 2
Summary of repositories

Partner	Repository name	URL	Type
UNIBO	**AMS Acta**	https://amsacta.unibo.it/	Institutional
UGOT	**Quality of Government (QoG)** **Zenodo**	http://qog.pol.gu.se/data https://zenodo.org/	Institutional Multi-disciplinary
IEA	**Zenodo**	https://zenodo.org/	
IAFE-NRI	**Zenodo**	https://zenodo.org/	Multi-disciplinary
UB	**Dipòsit Digital de la Universitat de Barcelona**	http://diposit.ub.edu	Institutional
PBS	**Portsmouth Research Portal** **Zenodo**	https://researchportal.port.ac.uk/portal/en/ https://zenodo.org/	Institutional Multi-disciplinary
WU	**Zenodo**	https://zenodo.org/	Multi-disciplinary

The table shows the repositories chosen by the partners for both dissemination and long term preservation.

As a general rule, *Zenodo* will be used for open dissemination and preservation of research data by all research teams that do not have suitable institutional, national, or disciplinary data repositories.

UGOT team will deposit its data sets both in *Zenodo* to comply with OpenAIRE requirements and in the portal of the Quality of Government Institute of the University of Gothenburg (QoG, see Table 2) for visibility and dissemination.

Portsmouth Research Portal is a CRIS system enabled to receive data attributing them valid DOIs but it is not currently indexed by OpenAIRE. Portsmouth University aims to get the repository indexed by OpenAire soon. If by the time of the publication of the project data, the repository will not be able to accept data sets yet, *Zenodo* will be used as dissemination and long term archiving solution by PBS.

To facilitate intelligibility and reuse, the data sets will be deposited in the data repositories along with all relevant documentation explaining data collection procedures and analysis.
In general, there will be no need to use specific software to access project data, since researchers will convert the data into open formats. In case particular software is used in data processing, full explanation and instructions will be included in the deposited documentation (a summary of the tools and software necessary to reuse of data sets is described in Table 3).

Table 3
Summary of tools and software for enabling re-use of the data sets

Tools/software
open spreadsheet and document editors, such as *OpenOffice*[12] or *LibreOffice*[13]
free CSV file viewers, such as *CSV viewer*[14]
R[15], free software environment for statistical computing and graphics
Mallet[16], open topic modelling software
open or free image viewers

A special case is represented by the model developed within WP6 to be processed through simulation in order to permit a scenario analysis: for the simulation code, a proprietary

[12] OpenOffice, http://www.openoffice.org/
[13] LibreOffice, https://www.libreoffice.org/
[14] CSV viewer, http://www.csvviewer.com/
[15] R, https://www.r-project.org/
[16] Mallet, http://mallet.cs.umass.edu/

software is needed ("*Vensim*"[17]). To make the simulation model available for reuse, the mathematical model will be saved using standard differential equations symbols in .rtf, .csv and .txt files. In this way the model can be simulated by any simulation software both open source and proprietary.

All data containing personal or sensible data will be shared only in anonymized form. In case of specific requests of access to restricted data by single researchers, research institutions, reviewers and committee, aimed for example at verifying the quality of the research results and at reproducing them, UNIBO will act as contact point and will evaluate each request consulting the Partner(s) that produced the requested data.

2.3 Making data interoperable

All data sets will be described using standard descriptive metadata, such as Dublin Core[7] and DataCite Metadata Schema[8] in order to ensure metadata interoperability for indexing and discoverability. All relevant documentation explaining codebooks, users' manuals, data collection procedures and analysis will be made available along with the data in order to guarantee intelligibility, reproducibility and the validation of the project findings.

Categories regarding official data relative to EU funded projects included in Annex II of commission regulation (EC) No 1828/2006[11] will be used to name the variables analyzed within the project. In the same way variable names of data derived from other official sources, such as Eurostat, will be consistent with the original source names. Variable names of data derived from surveys will match the survey question items as closely as possible.

2.4 Increase data re-use (licensing)

PERCEIVE will distribute the shareable data by adopting licenses that allow re-use of the data and of the data sets in their entirety by other scholars and stakeholders. The data sets will be made available mainly under Creative Commons license CC BY 4.0[18] and Open Data Commons ODC-BY[19].

CC BY 4.0 license permits users to freely share, modify, and use the data, subject only to full credit to the author(s). As an exception, CC BY NC 4.0[20], which requires full credit but limits reuse for commercial purposes, will be chosen when the data is collected from pre-existing sources that limit their free reuse (for example, when exception for illustration for scientific research is applicable, the reproduction of short excerpts will be possible only for not-commercial purposes). ODC-BY is a license specifically drafted for Open Data projects that works under condition of compatibility with Open Access requirements, interoperability and reuse.

(See the "Accessibility" section of the data sets descriptive tables in *Annex I* to check the corresponding license).

[17] Vensim (from Ventana System), https://vensim.com/
[18] Creative Commons Attribution (CC BY) 4.0 International, https://creativecommons.org/licenses/by/4.0/legalcode
[19] Open Data Commons Attribution License (ODC-By) v1.0, http://opendatacommons.org/licenses/by/1-0/
[20] Creative Commons Attribution-NonCommercial (CC BY NC) 4.0 International, https://creativecommons.org/licenses/by-nc/4.0/legalcode

In general, data will be made openly available as underlying data necessary to validate the research results immediately at the time of the publication of the corresponding scientific papers and public reports. Some data sets are expected to be part of public deliverables, in these cases data will be made available at the time of the release of the corresponding deliverable.

It is possible that an embargo period may be applied to some data sets to allow full exploitation of research results by the Partners. However, at the moment, it is not possible to predict all the data sets to which the embargo will be applied and its duration.

(See the "Accessibility" section of the data sets descriptive tables in *Annex I* to check the corresponding license and conditions).

Data will be given full citation from official project publications and web site and they will be made available in open formats through institutional or public data repositories compliant with OpenAIRE requirements[21] that guarantee long term preservation to the archived items, therefore they will be re-usable by third parties also after the end of the project (see Table 2 for the list of the chosen data repositories).

As the data collected or generated by the project are heterogeneous, the quality of the data will be carefully assured using different approaches. In WP1, the information relative to the EU financed projects will be resumed and harmonized using schemes designed specifically to collect and standardize information from different sources. The focus groups interviews, although held in local language and then translated in English, are based on a semi-structured questionnaire which assure an homogeneous harvesting of data. Data collection for the telephonic survey will be done via sub-contracting, and all available tools to generate a random, representative sample form each country will be employed. Moreover UGOT, which is the research team responsible for the survey, and the sub-contracting firm will generate design and population weights to better adjust the representativeness of the sample. In WP2, the quality control of UGOT data set will be done via a research assistant, checking the data against the original source. In WP3, the data will be normalized adjusting weights and assessing latent dimensions. In WP4, data relative to the urban policies for building smart cities will be normalized according to a specific procedure and robustness tests involving measurement errors will be carried out.[22] Moreover the variables constituting these data will be recorded using the following scheme: "Varname_year_source", e.g. employment_2015_ONS will represent data on employment with reference to year 2015 collected from ONS. In WP6, where a simulation model will be developed, a modelling team that mediates the relationship between modelers and other participants and facilitate the exchange of data, information and knowledge will be built up.

2.5 Allocation of resources

Making data FAIR requires a certain amount of researchers' time and investments in infrastructures. In PERCEIVE case, costs for long term deposit and preservation of public

[21] OpenAIRE, *For Data Providers* https://www.openaire.eu/intro-data-providers

[22] Greco, Salvatore and Ishizaka, Alessio and Matarazzo, Benedetto and Torrisi, Gianpiero (2015): Stochastic Multiattribute Acceptability Analysis: an application to the ranking of Italian regions. MPRA paper No. 75663. Munich.

shareable data are null because the chosen repositories do not apply fees for archiving and data curation.

During the project, a cloud storage solution will be adopted to share data among partners. The cost to activate and maintain it for the duration of the project will be covered by the project budget. The budget covers also the costs related to the project website setting up.

Costs related to data management and documentation, conversion of proprietary data files into open formats, and deposit procedures can be estimated about 3-5% of the amount of Person-Months assigned to each Partner for the research activities (from WP1 to WP6). A special case is represented by the time-consuming activities related to processing of interviews (i.e. transcription, translation and anonymization). Processing costs are estimated, for each Partner involved, about 0.1 Person-Months/hour of audio recording.

Moreover, the activities related to the DMP (such as providing guidance on data management and open access issues, coordinating the Partners, and preparing the DMP) will cost about 3 Person-Months for the whole duration of the project.

Responsible for data management are the data sets creators who are generally the team leaders directly involved in research data organization and collection (see Table 4).

Researchers are encouraged to identify themselves with the unique persistent identifier ORCID[23]. Registration is free of charge for researchers and allows for automated linkages between the researched identity and his research activities and outputs.

Table 4
Summary and contacts of the research team leaders responsible for the data sets

Team	Leader	ORCID ID (if available)	email
UNIBO	**Mollona, Edoardo**	http://orcid.org/0000-0001-9496-8618	edoardo.mollona@unibo.it
UGOT	**Charron, Nicholas**		nicholas.charron@pol.gu.se
IEA	**Tudor, Monica Mihaela**	http://orcid.org/0000-0002-8703-2296	monik_sena@yahoo.com
IAFE-NRI	**Chmieliński, Paweł**	http://orcid.org/0000-0001-8377-0702	Pawel.Chmielinski@ierigz.waw.pl
UB	**Suriñach Caralt, Jordi**	http://orcid.org/0000-0003-2622-3280	jsurinach@ub.edu
PBS	**Torrisi, Gianpiero**	http://orcid.org/0000-0003-4497-2365	gianpiero.torrisi@port.ac.uk
WU	**Meyer, Renate**		renate.meyer@wu.ac.at
	Höllerer, Markus		markus.hoellerer@wu.ac.at

Moreover, Partners are encouraged to identify and credit all contributors (see Table 5) participating in data management activities.

[23] ORCID, http://orcid.org/

Table 5
Summary of team members which contribute or are directly involved in the data sets creation and management

Team	Member	Role	ORCID ID (if available)
UNIBO	Brasili, Cristina	Project Member	http://orcid.org/0000-0003-2529-0158
	Pareschi, Luca	Researcher	http://orcid.org/0000-0002-4402-9329
	Reverberi, Pier Maurice	Researcher	
	Monasterolo, Irene	Researcher	
UGOT	Bauhr, Monika	Project Member	
IEA	Chitea, Mihai Alexandru	Researcher	http://orcid.org/0000-0001-5294-2604
	Florian, Violeta	Project Member	
	Rusu, Marioara	Project Member	
	Rosu, Elisabeta Stefania	Researcher	
	Chitea, Lorena Florentina	Researcher	
	Ionel, Iuliana	Researcher	
	Sima, Elena	Researcher	
IAFE-NRI	Wieliczko, Barbara	Project Member	
	Floriańczyk, Zbigniew	Project Member	
	Wigier, Marek	Project Member	
	Wasilewski, Adam	Project Member	
	Gospodarowicz, Marcin	Researcher	
UB	Royuela Mora, Vicente	Project Member	http://orcid.org/0000-0002-7647-0063
	López Bazo, Enrique	Project Member	http://orcid.org/0000-0002-4654-8237
	Moreno Serrano, Rosina	Project Member	http://orcid.org/0000-0002-3500-8622
WU	Barberio, Vitaliano	Project member	http://orcid.org/0000-0002-2615-5006
	Kuric, Ines	Researcher	
BAM!	Baruzzi, Davide	Project Member	
	Baruzzi, Lucia	Researcher	
	Spanò, Rosanna	Project Member	

Keys for "Role" column: Data Collector (such as survey conductors, interviewers…), Producer (person responsible for the form of a media product), Project Member (a researcher indicated in the GA), Researcher (an assistant to one of the authors who helped with research, data collection, processing and analysis but is not part of team indicated in the GA), Research Group (the name of a research institution or group that contributed to the data set).

(See *Annex I* for details about data management responsibilities related to each project data set).

2.6 Data security

Data shared among Partners will not contain sensitive data because they will be anonymized, with the only exception of data relative to interviewees who specifically ask not to be anonymized, such as public stakeholders.

At each institution, research data will be stored in computers, laptops, intranets or hard-drives accessible through institutional password periodically modified according to national law provisions for data security and protected by regularly updated antiviruses. None of the project data will be left inadvertently available.

All the research materials stored in computers are subject to regular backup in order to safeguard them from accidental losses. For example, thanks to a partnership with Google® offering unlimited storage space to the University of Portsmouth staff, PBS team will perform a secure backup of its data on Google Drive® on a daily base for the whole length of the project. All the data will be password protected. If mobile devices are used to store data files (e.g.

backup files), they will be kept in a safe place accessible only to the researchers involved or will be encrypted with *ad-hoc* software.

A cloud storage solution will be adopted for data sharing among research teams. In this case, as well, regular backup of the data will be performed to ensure data recovery. In addition all Partners are asked to keep local updated copies of all their files.

Long term preservation of public data is ensured by the chosen data repositories that have specific preservation policies.
UNIBO *AMS Acta* guarantees long term preservation to the archived materials also thanks to a deposit agreement with the National Central Library in Florence.
Data deposited in *Portsmouth Research Portal* will be preserved for at least 10 years as indicated in the university policy[24]. After this 10 year retention period, further retention will be reviewed by the repository managers.
UB *Dipòsit Digital de la Universitat de Barcelona* has a long term preservation policy without retention period limits.
Zenodo policy[25] ensures that the items will be retained for the lifetime of the repository and in case of closure, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.

2.7 Ethical aspects

Research in PERCEIVE involves questionnaires, interviews and surveys with adults participants and focus groups with key informants (selected practitioners and experts of the EU Cohesion Policy) . All aspects of collection of data involving personal data (in particular in WP1, WP2, WP3, WP4, WP6) are covered by the Ethics Requirements document (WP9).

All personal data collected within the PERCEIVE project from questionnaires, interviewers, surveys and focus groups are carefully protected in compliance with relevant national data protection legislation of the EU member states implementing the European directive 95/46/EC and with the procedures defined by the European Code of Conduct for Research Integrity.

As a general principle, personal data resulting from the focus groups, interviews, observation and questionnaires will be separated from the research results, and will be handled by different members of the research team. In regards to the respondents in the survey, they will be selected at random and their name and address will not be recorded. The data will be stored in a way not to allow the identification of the subject, adopting measures for anonymization (i.e. names replaced by initials or pseudonyms); results of questionnaires and interviews will be transmitted or made available to the other project partners as anonymous data.

In principle, research will not involve personal sensitive data. Notwithstanding, in certain circumstances, data collected during interviews and focus groups may potentially be sensitive (e.g. participants might disclose political opinions). Only personal characteristics that are strictly necessary, for theoretical reasons and to the benefit of the research, will be collected;

[24] Portsmouth Research Portal policy, http://www.port.ac.uk/library/help/research/researchdata/
[25] Zenodo policy, https://zenodo.org/policies

such characteristics may be used to compare participants and it will be clarified how such characteristics help to shape their views and opinions.

Files containing questionnaire data for statistical analysis, transcripts of interviews and focus groups, transcripts of field observations, photos, minutes, videos, action diaries, etc.) are stored in computers, laptops, intranets or hard-drives of the research institutions accessible through institutional password modified periodically (every 3 months in case of storage of sensitive data), and protected by regularly updated antiviruses. Files containing "sensitive" data will be stored encrypted. Password-protected and encrypted files are accessible only to authorized members of the research teams receiving preliminarily specific information and training on the procedures for data collection, storage etc. None of the project data will be left inadvertently available by being left on desks or in unlocked rooms. All the research materials stored in computers are subjected to back up regularly (according to each institutions' regulations) in order to safeguard them from accidental losses.

Data and information collected from questionnaires will be disseminated and published only in an aggregate and/or anonymous form. Publications will only report aggregate data and shall not contain information that may permit the identification of individual participants.

Data that are not shareable will be stored for the time required by the international scientific community (at least 5 years after the conclusion of the research project) and will be subsequently destroyed. Where personal data are no more necessary for the research, they will be immediately destroyed. Qualitative data files can be accessible with public access as long as any information that can lead to identification of an individual participant is deleted.

3. Data sets overview

The following table (Table 6) offers an overview of the data sets expected from the project and described more in detail in *Annex I*. It will be updated according to DMP changes and variations.

Table 6
Data sets list

Table acronyms and abbreviations: # = data set progressive number ID, LB = WP lead beneficiary, PP = project phase (starting month-ending month), CT = creator team in charge of curating the data set, C=collected, G=generated, A=available, IP=in progress, NYA=not yet available.

#	WP	LB	TASK	PP	CT	DATA SET Title	SOURCE	STATUS
1	1	UNIBO	1.1	1-9	**UNIBO**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.1: The framework for the comparative analysis***	**C,G**	**IP**
2	1	UNIBO	1.2	6-9	**UNIBO**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.2: Focus group with Cohesion Policy practitioners***	**G**	**NYA**
3	1	UNIBO	1.3	8-12	**UGOT**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.3: Survey at citizen level***	**G**	**IP**
4	2	UNIBO	2.1	1-12	**UGOT**	***PERCEIVE: WP2: Evaluation of EU citizens' perception of the EU project: Task2.1: Data relative to regional performance of the Cohesion Policy and institutional quality***	**C,G**	**IP**
5	2	UB	2.2 2.3 2.4 2.5	12-28	**UB**	***PERCEIVE: WP2: Evaluation of EU citizens' perception of the EU project: Key variables for citizens' identification with EU***	**C,G**	**NYA**
6	3	WU	3.1 3.2 3.3 3.4 3.5	6-22	**WU**	***PERCEIVE: WP3: Effectiveness of communication strategies of EU projects***	**C,G**	**NYA**
7	4	PBS	4.1 4.2 4.3	1-12	**PBS**	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: City smartness***	**C,G**	**IP**
8	4	PBS	4.4	12-14	**IAFE-NRI**	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: Task4.4: Cohesion Policy vs Urban and Rural policies to address spatial discrepancies in EU territorial policy***	**G**	**NYA**
9	4	PBS	4.5	13-15	**IEA**	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: Task4.5: Citizens' vs Practitioners' views of the EU Cohesion Policy***	**G**	**NYA**
10	5	WU	5.1 5.2 5.3 5.4 5.5	1-17	**WU**	***PERCEIVE: WP5: The multiplicity of shared meanings of EU and Cohesion Regional and Urban Policy at different discursive levels***	**C,G**	**NYA**
11	6	UNIBO	6.1 6.2 6.3 6.4	22-36	**UNIBO**	***PERCEIVE: WP6: Virtual learning environment: Simulation and modelling data***	**G**	**NYA**

Annex I: Data sets tables

The analytic descriptions of the expected data sets of PERCEIVE project are reported in this Annex and organized by work-packages.

WP1 - Framework for comparative analysis: differences, implementation, perceptions and data gathering

WP1 will set the framework for the theoretical development for the next stages of the project research activities. In particular, the multiplicity and logical organization of citizens and practitioners' understandings of Cohesion Policy and its implementation (citizens' European/national/regional identity nesting and identification) will be analyzed by means of a combination of different quantitative and qualitative methods.

Lead: **UNIBO**

Participants: **UGOT, IEA, IAFE-NRI, UB, PBS, WU, BAM!**

Months: **1-12**

1	**In progress**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.1: The framework for the comparative analysis***
ID [ID type]		
Team in charge		UNIBO
Creator/s		Brasili, Cristina [UNIBO]
Contributor/s		Charron, Nicholas [UGOT], Tudor, Monica Mihaela [IEA], Chmieliński, Paweł [IAFE-NRI], Suriñach Caralt, Jordi [UB], Torrisi, Gianpiero [PBS], Barberio, Vitaliano [WU], Meyer, Renate [WU], Höllerer, Markus [WU]
Contact Person/s		Reverberi, Pierre Maurice [UNIBO, pierre.reverberi2@unibo.it]
Contents		This dataset contains data on the Cohesion Policy projects managed by Local Managing Authorities in the selected case-studies, ie: Emilia-Romagna and Calabria (IT), Extremadura (ES), Burgenland (AT), Norra Mellansverige (SE), Sud-Est (RO), Essex (UK), Dolnośląskie and Warmińsko-Mazurskie (PL), in the programming period 2007-2013. Information includes objectives of the projects, the projects' funding broken down by source and year, territorial coverage, authorities involved, as well as information on beneficiaries. Most of data were already produced and transmitted to the European Commission for the monitoring of the Cohesion Policy projects and published, in aggregate form, on European and national portals. Each partner contributes by gathering data referred to the selected case studies in the partner's Country in the disaggregated form. The dataset is made of 4 sheets, plus one with definitions, collecting information on: (1.) scope and location of the projects, as categorized by the European Commission in the Annex II of Commission Regulation (EC) No 1828/2006 (p. 53); (2.) payments: the amount of funding received by CP projects broken down by source and year; (3.) beneficiaries; (4.) expenditure for communication activities. Data in sheets 1 and 2 were already

1	**In progress**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.1: The framework for the comparative analysis***
		produced by the Local Managing Authorities and transmitted to the European Commission for the monitoring of the Cohesion Policy projects. Sheets 3 and 4 were designed to collect and standardize information on beneficiaries of the projects funded with Structural Funds and project communication plans. The following Local Managing Authorities are involved in PERCEIVE as third-parties, and they directly contribute to the dataset: Regione Emilia-Romagna (IT), Regione Calabria (IT), Ministerio de Economìa y Hacienda - Dirección General de Fondos Comunitarios (ES) and Agentia pentru Dezvoltare Regionala Sud-Est (RO). Potential users of the dataset include researchers, policymakers and stakeholders of Cohesion Policy.
Data format		Qualitative and quantitative tabular data (.csv).
Data Volume		Final volume of data is expected to be 50 MB.
Accessibility		The dataset will be made available as part of deliverable D1.1 "Report with regional case-studies description and analysis of the focus groups with third-parties' Cohesion Policy practitioners and stakeholders and database of the Cohesion Policy projects managed by case-studies' Local Managing Authorities in the programming period 2007-2013. It will be available under a CC BY 4.0 license.

2	**Not yet available**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.2: Focus group with Cohesion Policy practitioners***
ID [ID type]		
Team in charge		UNIBO
Creator/s		Brasili, Cristina [UNIBO]
Contributor/s		Pareschi, Luca [UNIBO], Charron, Nicholas [UGOT], Tudor, Monica Mihaela [IEA], Chmieliński, Paweł [IAFE-NRI], Suriñach Caralt, Jordi [UB], Torrisi, Gianpiero [PBS], Barberio, Vitaliano [WU], Meyer, Renate [WU], Höllerer, Markus [WU]
Contact Person/s		Reverberi, Pierre Maurice [UNIBO, pierre.reverberi2@unibo.it]
Contents		The data set consists of transcripts of interviews to practitioners, beneficiaries and targets of the Cohesion Policy projects in the selected case-study regions. Interviewees are asked to provide their views and perceptions on the multilevel governance system, on the communication activities of the Operational Programmes and on the effectiveness of Cohesion Policy. Interviews are held at the Local Managing Authorities of the case-study regions of the PERCEIVE project: Emilia-Romagna and Calabria (IT), Extremadura (ES), Burgenland (AT), Norra Mellansverige (SE), Sud-Est (RO), Essex (UK), Dolnośląskie and Warmińsko-Mazurskie (PL). The focus groups are based on a semi-structured questionnaire. Each focus group is held in the local language and recorded with audio equipment. Each partner is then in charge to produce interview transcripts and to translate them in English.
Data format		Qualitative textual data (.txt).

2	**Not yet available**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.2: Focus group with Cohesion Policy practitioners***
Data Volume		Final volume of data is expected to be 20 MB.
Accessibility		Raw data (audio recordings of focus group) will not be preserved or shared in compliance with the project's ethical rules. Interviewees will be asked for their consent on the usage of the interviews. For those participants who will opt for anonymization of the data, this will be ensured when producing the transcript. Audio recordings will then be deleted and only the processed transcripts will be shared and disseminated. Shareable data will be released as part of Deliverable 1.1 " Report with regional case-studies description and analysis of the focus groups with third-parties' Cohesion Policy practitioners and stakeholders and database of the Cohesion Policy projects managed by case-studies' Local Managing Authorities in the programming period 2007-2013 (month 9) but they are expected also as underlying data necessary to validate the results presented in scientific publications, immediately at publication time under a CC BY 4.0 license.

3	**In progress**	***PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.3: Survey at citizen level***
ID [ID type]		
Team in charge		UGOT
Creator/s		Charron, Nicholas [UGOT]
Contributor/s		Bauhr, Monika [UGOT]
Contact Person/s		Charron, Nicholas [UGOT, Nicholas.charron@pol.gu.se]
Contents		The data set contains numerical and textual tabular data converted into digital format (survey made by telephone), quantitative and qualitative data, raw or processed. The survey will involve 20.000 interviewees in the countries selected for case studies (Italy, Spain, Sweden, United Kingdom, Romania, Poland, Austria) and in control countries (Netherlands, Slovakia, Latvia, Germany, France). Data will be micro-level with regional and country level identifiers. The goal will be to compare support and awareness of EU Cohesion Policy across regions and countries using individual level explanations derived from survey indicators.
Data format		Statistical tabular data (.csv and STATA format: .dta)
Data Volume		Final volume of data is expected to be 24 MB.
Accessibility		No personal data will be collected, and thus all data can be shared under a ODC license

WP2 - Evaluation of EU citizens' perception of the EU project in relation to regional performance of the Cohesion Policy and institutional quality

WP2 will focus on understanding the relation between the EU Cohesion Policy performance and citizens' perception of it, accounting for the influential role of institutional quality and heterogeneity of EU integration paths and experiences in different EU regions.
In particular, it will provide a quantitative representation (in terms of variables and indexes) of the key concepts at the backbone of the project objectives and rationale; it will analyze the quantitative relation/functional form behind the socio economic / cultural / geographical / political multidimensional determinants of EU citizens' identity; and it will use alternative quantitative methodologies to analyze the evolution of citizens' identification with the EU project in different territories, and the co-evolution with key socio-economic magnitudes such as income per capita, unemployment, and EU funds allocated in each territory.

Lead: **UB**

Participants: **UNIBO, UGOT, IEA, IAFE-NRI, PBS, WU**

Months: **1-32**

4	**In progress**	***PERCEIVE: WP2: Evaluation of EU citizens' perception of the EU project: Task2.1: Data relative to regional performance of the Cohesion Policy and institutional quality***
ID [ID type]		
Team in charge		UGOT
Creator/s		Charron, Nicholas [UGOT]
Contributor/s		UNIBO, IEA, IAFE-NRI, UB, PBS, WU
Contact Person/s		Charron, Nicholas [UGOT, Nicholas.charron@pol.gu.se]
Contents		This data set will include different primary data: European projects data from *"PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.1: The framework for the comparative analysis"* data set, focus groups data from *"PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.2: Focus group with Cohesion Policy practitioners"* data set, and survey data from *"PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.3: Survey at citizen level"* data set; as well as secondary data gathered from other existing sources (Eurobarometer, Eurostat). The unit of analysis will be the region (NUTS[26] 1, 2, and 3 will be specified).
Data format		Statistical tabular data (.csv and STATA format: .dta).
Data Volume		Final volume of data is expected to be 6 MB.

[26] Nomenclature of territorial units for statistics (NUTS), more information about the current NUTS 2013 classification at http://ec.europa.eu/eurostat/web/nuts/overview

4	**In progress**	***PERCEIVE: WP2: Evaluation of EU citizens' perception of the EU project: Task2.1: Data relative to regional performance of the Cohesion Policy and institutional quality***
Accessibility		After gaining copyright permission from original sources the data set will be made openly available under a ODC – BY license. It is not defined yet whether the data set will be made openly available as underlying data at the time of publication of research results or as other data.

5	**Not yet available**	***PERCEIVE: WP2: Evaluation of EU citizens' perception of the EU project: Key variables for citizens' identification with EU***
ID [ID type]		[Handle]
Team in charge		UB
Creator/s		Royuela, Vicente [UB]
Contributor/s		Suriñach, Jordi [UB]; López Bazo , Enrique [UB]; Moreno, Rosina [UB]; Monasterolo, Irene [UNIBO]; [UNIBO], [IEA], [IAFE-NRI], [UB], [PBS], [WU]
Contact Person/s		Royuela, Vicente [UB, vroyuela@ub.edu]
Contents		This dataset refers to PERCEIVE tasks 2.2, 2.3, 2.4 and 2.5. These tasks analyze the data sets developed by PERCEIVE task 2.1 in order to identify, quantify and synthetize within one functional form (mathematical formula) the most important variables to asses identification with EU by citizens. The results will feed an update of the data set developed in task 2.1. Finally, these data will be used to model convergence of EU citizens' identification in the case study regions. Data will consist of tables, results of cluster analysis and validation of the mathematical formula. Moreover, data on absorption of EU funds, provided by the EC and other research centres and collected in task 2.5, will be gathered in this data set, together with the results of convergence analysis. This data set will be used to evaluate the EU citizens' perception of the EU project in relation to regional performance of the cohesion policy and institutional quality. WP2 will use primary data from the survey developed in WP1 (*"PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.3: Survey at citizen level"* data set) and also comparative case studies on EU countries and regions, coming also from WP1 (*"PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.1: The framework for the comparative analysis"* and *"PERCEIVE: WP1: Framework for comparative analysis of the perception of Cohesion Policy and identification with the European Union at citizen level in different European countries: Task1.2: Focus group with Cohesion Policy practitioners" data sets*). The data set will incorporate data from existing sources including Eurostat and information resulting from the Survey conducted at WP1 of the PERCEIVE's project. As regards secondary data, no additional treatment is needed. Info about the sources of such data, including references to technical details, will be provided in the dataset's codebook. A tentative list of topics to be extracted from Eurostat: - Regional economic accounts ESA 2010/2005, Gross domestic product indicators - Regional science and technology statistics, R&D expenditure and personnel - Regional science and technology statistics, Employment in high-tech sectors - Regional science and technology statistics, Intellectual property rights, Patent

5	**Not yet available**	***PERCEIVE: WP2: Evaluation of EU citizens' perception of the EU project: Key variables for citizens' identification with EU***
		- Regional labour market statistics, Regional employment - LFS annual series - Regional labour market statistics, Regional employment - LFS annual series, Educational attainment - Regional transport statistics, Road and rail networks
Data format		Numerical and textual data (.csv).
Accessibility		For information resulting from PERCEIVE WP1 the data set will follow the same protection as described in the DMP of this WP. For the rest of the information, as it will be resulting from open access sources, such as Eurostat, there will be no protection. Data will be openly shared at the time of publication of results under a CC BY NC 4.0 license.

WP3 - The effectiveness of projects' communication strategies

WP3 will be focused on the study of the communication strategies of EU financed projects.
In particular, it will map material and rhetorical strategies that are employed in the communication efforts of a variety of EU financed projects to build legitimacy. In addition, the ways in which the content of several genres of media are generated centrally and locally adapted will be measured with a specific focus on symbolic aspects of communication; the impact and the effectiveness of communication efforts related to EU-financed regional development projects will be both assessed in terms of variations in the level of citizens' awareness and appreciation of the projects themselves; and the historical-dynamic dimension of communication efforts concerning EU financed projects will be reconstructed.

Lead: **WU**

Participants: **UNIBO, UGOT, IEA, IAFE-NRI, UB, PBS, BAM!**

Months: **6-22**

6	**Not yet available**	***PERCEIVE: WP3: Effectiveness of communication strategies of EU projects***
ID [ID type]		
Team in charge		WU
Creator/s		Barberio, Vitaliano [WU];
Contributor/s		Pareschi, Luca [UNIBO]; Mollona, Edoardo [UNIBO]; Meyer, Renate [WU], Höllerer, Markus [WU]
Contact Person/s		Barberio, Vitaliano [WU,]
Contents		This data set will contain all the relevant data referred to PERCEIVE WP3. Indeed the tasks within WP3, despite consisting of a number of different ones, are conceptually one stream of research only due to the project's structure. In particular, we will collect information from interviews and questionnaires (Task 3.1), a comprehensive list of projects and their managers (Task 3.2), data collected at national and European level regarding communication from EU communication offices (Task 3.3), the results of the analysis of these communications (Task 3.4), and data extracted from instruments such as Eurobarometer in order to have a longitudinal understanding of the identification and awareness (Task 3.5). Textual and numerical data will be included in this data set. To our knowledge, there is no comparable public source of data. Therefore these data will be helpful for all students, researcher, and practitioners willing to understand how to effectively and efficiently communicate policy.
Data format		Numerical and textual data (.csv, .rtf).
Data volume		Final volume of data is expected to be 500 MB.
Accessibility		Results of the survey will be anonymized and clustered at NUTS2 level. Conversion into standard open formats will be performed. When using copyrighted materials, proper citation of the source will be made available. Names and email addresses of survey recipients will be neither recorded nor shared. To ensure long term preservation data will be anonymized and documented in attached read-me files. It is not defined yet whether the data set will be made openly available as underlying data at the time of publication of research results or as other data.

6	**Not yet available**	***PERCEIVE: WP3: Effectiveness of communication strategies of EU projects***
		The data set will be released under a CC BY NC 4.0 license.

WP4 - Spatial determinants of policy performance and synergies

WP4 will focus on better understanding the spatial discrepancies of the EU citizens' perceptions of European Policy performance, in order to improve them.
In particular, the identification of possible complementarities and synergies between EU Cohesion Policy, rural and urban policies supported by EU funds will be performed; an innovative index of city's "smartness" will be defined exploring the interplay between EU Urban Policy and the implementation of "smart cities"; and the discrepancies between regional policies implemented in each case study area and the "real problems" perceived by regional actors / population will be in-depth analysed.

Lead: **PBS**

Participants: **UNIBO, IEA, IAFE-NRI, UB, WU**

Months: **1-15**

7	**In progress**	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: City smartness***
ID [ID type]		
Team in charge		PBS
Creator/s		Torrisi, Gianpiero (PBS)
Contributor/s		
Contact Person/s		Torrisi, Gianpiero [PBS, Gianpiero.torrisi@port.ac.uk]
Contents		The dataset contains data on both smart cities projects and smart cities characteristics to be used for the computation of the composite indicators using the Stochastic Multi-criteria Acceptability Analysis (SMAA) methodology. The dataset reports the data used to analyze the concept of a 'smart city' along two main dimensions. First, it contains the data to operationalize the concept of smart city along the dimensions elaborated in the ongoing literature including proxies for networked infrastructure to improve economic and political efficiency and enable social and cultural development, the extent of business-led development, the social inclusion of various urban residents in public services, extent of high-tech and creative industries, social and relational capital, and social and environmental sustainability. Second, it reports the results of the analysis using the above dimensions to compute a new index of smartness and quality of life based on SMAA. The data set will include secondary data collected from official sources such as Eurostat and Office for National Statistics (ONS). More in detail, the dataset uses EU data on cities and functional urban area from the Urban Audit dataset (http://ec.europa.eu/eurostat/web/cities). These EU data are joined to ONS data belonging to the Towns and Cities Analysis database (https://www.ons.gov.uk/releases/townsandcitiesanalysis). All the data will be normalized according the procedure described in Greco et al. (2015). Furthermore, robustness tests involving measurement errors will be carried out according to the procedure introduced in the aforementioned paper by Greco et al. (2015).
Data format		Numerical tabular data (.csv).
Accessibility		Not yet defined whether the data set will be made openly available as underlying data or other data. It will be made available under a CC BY 4.0 license.

8	Not yet available	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: Task4.4: Cohesion Policy vs Urban and Rural policies to address spatial discrepancies in EU territorial policy***
ID [ID type]		
Team in charge		IAFE-NRI
Creator/s		Chmieliński, Paweł [IAFE-NRI]
Contributor/s		Wieliczko, Barbara [IAFE-NRI]; Floriańczyk, Zbigniew [IAFE-NRI]; Wigier, Marek [IAFE-NRI]; Wasilewski, Adam [IAFE-NRI]; Gospodarowicz, Marcin [IAFE-NRI]
Contact Person/s		Chmieliński Paweł [IAFE-NRI, Pawel.Chmielinski@ierigz.waw.pl]
Contents		This dataset will collect and generate data addressing the relationship between territorial cohesion objectives and the problems perceived by citizens. In particular a comparative analysis between the case study regions will generate data useful for identifying best practices in mixing the EU policy instruments for a better achievement of regional needs. Data that will be generated via focus groups interviews among 5-10 representatives of LMA (local management authorities) in 2 Polish regions: Dolnośląskie and Warmińsko-Mazurskie . Interview transcripts will be use to address how territorial cohesion objectives match the “real problems” of regions. The focus groups are built around the following main topics: 1) governance of the Cohesion Policy projects, in order to understand how different authorities at different levels cooperate and share the responsibilities for the implementation of the Cohesion Policy; 2) level of citizen engagement, in order to understand whether a bottom-up approach is used; 3) how the media inform on the Cohesion Policy programmes, in order to appreciate the discrepancies (if any) about the aims of Cohesion Policy and its construction on the public discourse. Comparing current and past programming periods, we will investigate how the policy performs in reducing the gap between territorial cohesion objectives and “real problems” defined by LMAs and citizens. Because it is project-specific data and reflects the concept and methodology of the study under PERCEIVE they are perceived as unique - similar data does not exist. Potential users will be Regional Policy’s European/National/Local policy makers and practitioners, European networks and associations looking to data on LMA opinions on cohesion policy implementation in Poland to be used in policy recommendation, studies and policy making process; next group of potential data users are researchers working on assessment of Cohesion Policy, data may be used as a source for topic-related studies, case studies, comparisons.
Data format		Qualitative textual data (.rtf).
Accessibility		Personal data will be kept confidential due to personal data protection laws and ethical reasons. Anonymized data will be shared but it is not defined yet whether they will be made available as underlying data or as other data. It will be made available under a CC BY 4.0 license.

9	Not yet available	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: Task4.5: Citizens' vs Practitioners' views of the EU Cohesion Policy***
ID [ID type]		
Team in charge		IEA
Creator/s		Tudor, Monica Mihaela [IEA]

9	Not yet available	***PERCEIVE: WP4: Spatial determinants of policy performance and synergies: Task4.5: Citizens' vs Practitioners' views of the EU Cohesion Policy***
Contributor/s		Florian, Violeta [IEA]; Rusu, Marioara [IEA]; Rosu, Elisabeta Stefania [IEA]; Chitea, Lorena Florentina [IEA]; Ionel, Iuliana [IEA]; Sima, Elena [IEA]
Contact Person/s		Chitea, Mihai Alexandru [IEA, mihai_chitea@yahoo.com]
Contents		The data set will include data generated by PERCEIVE project Task 4.5. based on the results of the focus group and of the survey performed in PERCEIVE WP1 (hierarchy of problems for citizens and objectives of Cohesion Policies). A hierarchy / evaluation scale will be conceptualized for the identified discrepancies, which will be applied for each case study area. The analysis will use the contingency method to compare the European Cohesion Policy objectives, on one hand, and the hierarchy of real problems identified by the regional actors obtained though the focus group method. The correlation analysis will be used to analyze the intensity of the relation between the size of previously identified discrepancies and the perception on the cohesion policy efficiency. The data set will allow to understand whether the EU Cohesion Policy is perceived and understood by the citizens in the same way as it is conceived by practitioners. Potential users include researchers, students, policy makers as well as others that can use the data for different analyses.
Data format		Quantitative numerical data (.csv) and qualitative textual data (.rtf).
Data volume		Final volume of data is expected to be 20 MB.
Accessibility		It is not defined yet whether the data set will be made available as underlying data or as other data. The data set will be released under a ODC – BY license

WP5 - The multiplicity of shared meanings of EU and Cohesion Regional and Urban Policy at different discursive levels

WP5 will focus on filling the current knowledge gap about the cultural-cognitive multiplicity of meanings (logics of) Cohesion Policy and European identity and identification.
In particular, an original conceptual account for the importance of social construction, especially discourse, in understanding the impact that cohesion policy might have on EU identity and identification will be advanced. Discourse about EU implementation of Cohesion Policy and European identity will be explored both qualitatively and quantitatively, disentangling several elements of meanings multiplicity and complexity, such as topics and their composition and internal logics. Moreover, the diverse contribution of relevant actors (such as the communication organs of the EU, the local press and activists, …) "theorizing" the EU Cohesion Policy and identity at different levels will be assessed. Then, interregional variations in the association between the presence, co-presence and relative intensity of multiple qualified topics in local discourses on the one hand, and citizens' selection of an "appropriate" logic of European identity and identification with the EU on the other hand will be assessed in order to validate the proposed conceptual account.

Lead: **WU**

Participants: **UNIBO, UGOT, IEA, IAFE-NRI, UB, PBS**

Months: **1-17**

10	**Not yet available**	***PERCEIVE: WP5: The multiplicity of shared meanings of EU and Cohesion Regional and Urban Policy at different discursive levels***
ID [ID type]		
Team in charge		WU
Creator/s		Barberio, Vitaliano [WU],
Contributor/s		Pareschi, Luca [UNIBO], Mollona, Edoardo [UNIBO], Meyer, Renate [WU], Höllerer, Markus [WU]
Contact Person/s		Barberio, Vitaliano [WU, vitaliano.barberio@wu.ac.at]
Contents		The data set will include a large collection of data from different documentary sources and media: EU policies and reports, descriptions and reports created by Local Managing Authorities, newspaper articles, tweets, blog posts referred to EU CP policies. The data set will also include analysis performed on data: we will use Mallet, an open source software aimed at creating topics, that are sets of words that co-occur together in a meaningful way. The results of the linear regression analysis will be in this data set as well. The results of the analysis will consist of tables of texts and of numerical data. Collected data are only partly available online, therefore our generated data will have a unique value, as there is no comparable public source of data. Data will be helpful for all student and practitioners willing to understand how the concepts of Cohesion Policies, Europe and European identity are shaped in the public sphere.
Data format		Numerical and textual tabular and qualitative data (.csv, .rtf, .txt).
Data volume		Final volume of data is expected to be 2 GB.
Accessibility		Raw data consisting in collected copyrighted materials shall not be made available

10	**Not yet available**	***PERCEIVE: WP5: The multiplicity of shared meanings of EU and Cohesion Regional and Urban Policy at different discursive levels***
		without permission of the right holders. Personal data derived from social networks will be anonymized. CC BY NC 4.0 (this license will be used for citation of excerpts of copyrighted materials). CC BY 4.0 (this license will be used in all other cases)

WP6 - Virtual learning environment

WP6 is aimed at integrating contributions from the different PERCEIVE research areas and to generate a model for the analysis of the interaction among Cohesion Policy implementation, regional specificities, communication policies, the symbolic space that is created by regional policies and identification with the EU project. The model will make available a virtual learning environment for policy-makers. By employing simulation analysis, policy-makers will be able to perform scenario analysis to explore consequences of communication strategies and of cohesion policy implementation.

Lead: **UNIBO**

Participants: **IEA, IAFE-NRI, UB, PBS, WU, BAM!**

Months: **22-36**

11	**Not yet available**	***PERCEIVE: WP6: Virtual learning environment: simulation and modelling data***
ID [ID type]		
Team in charge		UNIBO
Creator/s		Mollona, Edoardo [UNIBO]
Contributor/s		Pareschi, Luca [UNIBO], [BAM!]
Contact Person/s		Mollona, Edoardo [UNIBO, edoardo.mollona@unibo.it]
Contents		The data set will include data generated by PERCEIVE WP6 (Tasks 6.1, 6.2, 6.3 and 6.4). These tasks refer to simulation, which entails the building and the validation of both a qualitative and quantitative model and of a number of simulation experiments. The nature of the data is mixed. First, the data set will include the strings of codes that are necessary to simulate the mathematical model. Second, the data set includes the numerical parameters necessary to produce specific simulation experiments. Third, the data set makes available the simulation results reported as simulated time series. As for their nature, the data will be the outcome of an original modelling effort. The researchers will transform into a mathematical model the knowledge extracted from different sources: data generated by focus groups, data flowing from the survey (PERCEIVE WP1), data extracted from a set of relevant contributions available in the academic and practitioner literature (the set is not yet available) and the workshops in which the simulation model will be tested. Potential users include researchers, students, policy makers as well as others that can use the data to simulate and explore the scenario that follow from alternative specific communication policies.
Data format		Numerical and textual data (.csv, .rtf, and .txt). NB: for simulation code, a proprietary software is needed. To make the simulation model eventually available for reuse, we will save the mathematical model using standard differential equations symbols into file .rtf, .csv and .txt files. As a matter of fact, the model can be simulated (reused) using various simulation software (open or proprietary), provided that the mathematical model is available.
Data volume		Final volume of data is expected to be 100 MB.

11	**Not yet available**	***PERCEIVE: WP6: Virtual learning environment: simulation and modelling data***
Accessibility		At the moment it is not defined whether the data set will be made available as underlying data at the time of publication of results or as other data. The data set will be released under a CC BY 4.0 license.