i REACT

IMPROVING RESILIENCE TO EMERGENCIES THROUGH

ADVANCED CYBER TECHNOLOGIES

Data Management Plan

Deliverable ID	D2.8
Work Package Reference	WP2
Issue	1.0
Due Date of Deliverable	30/11/2016
Submission Date	02/12/2016
Dissemination Level[1]	PU
Lead Partner	POLITO
Contributors	POLITO, ISMB
Grant Agreement No	700256
Call ID	H2020-DRS-1-2015
Funding Scheme	Collaborative

European Commission

I-REACT is co-founded by the Horizon 2020 Framework Programme of the European Commission under grant agreement n. 700256

Prepared by	Reviewed by	Approved by
P. Garza	C.Rossi	F. Dominici

Issue	Date	Description	Author(s)
1.0	29/11/2016	First issue of the deliverable	P. Garza
0.1	1/11/2016	First draft of the deliverable	P. Garza

Project: I-REACT
Grant Agreement: 700256
Data Management Plan
Call ID: H2020-DRS-1-2015
Deliverable ID: D2.8
Page: 2 of 14

TABLE OF CONTENTS

Project: I-REACT

Grant Agreement: 700256

Data Management Plan

Call ID: H2020-DRS-1-2015

Deliverable ID: D2.8

Page: 3 of 14

1 INTRODUCTION

1.1 PURPOSE OF THE DOCUMENT

This document describes how we defined the Data Management Plan (DMP) of I-REACT and contains its first version. The DMP covers the data management life cycle of the collected, processed and generated data during the project. Specifically, the DMP is used to define the guidelines for data management in the project in order to ensure a high level of data quality, security, and accessibility.

According to the "Guidelines on FAIR Data Management in Horizon 2020" [RD01], "*the DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur.*".

This deliverable contains the first version of the DMP. The DMP will be updated, providing finer details, during the periodic evaluation/assessment of the project (i.e., at the mid-term review and at the final review).

We prepared the Data Management Plan by following the template provided in the "Guidelines on FAIR Data Management in Horizon 2020" [RD01] document. Specifically, we generated the DMP of I-REACT by using the DMP online [RD02] tool, which is compatible and compliant with the requirements set out in Annex 1 of the "Guidelines on FAIR Data Management in Horizon 2020" [RD01]. DMP online supports the generation of the DMP and allows exporting it in different electronic formats (e.g., pdf file, doc file).

The DMP of I-REACT follows the FAIR principles, i.e., research data are made Findable, Accessible, Interoperable and Re-usable.

1.2 STRUCTURE OF THE DOCUMENT

The document is organized as it follows:

- **Chapter 1** is this introduction and description of the document itself;
- **Chapter 2** contains the Data Management Plan of I-REACT. Specifically, the content of Chapter 2 corresponds to the export of the Data Management Plan (DMP) that we created by means of the DMP online [RD02] tool;
- **Chapter 3** draws conclusions.

Project: I-REACT
Grant Agreement: 700256

Data Management Plan
Call ID: H2020-DRS-1-2015

Deliverable ID: D2.8
Page: 4 of 14

1.3 ACRONYMS LIST

DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Re-usable

1.4 REFERENCE AND APPLICABLE DOCUMENTS

ID	Title	Revision	Date
[RD01]	EC's Guidelines on Data Management in Horizon 2020	3.0	26/07/2016
[RD02]	DMP online https://dmponline.dcc.ac.uk	-	
[RD03]	Zenodo repository https://www.zenodo.org		
[RD04]	D2.4 Report on Privacy and Security		

2 I-REACT: DATA MANAGEMENT PLAN (DMP V1.0)

This section contains the DMP that has been generated by using the DMP online [RD02] tool. For the generation of the DMP we selected the H2020 template provided by the DMP online tool, which is compatible and compliant with the requirements set out in Annex 1 of the "Guidelines on FAIR Data Management in Horizon 2020" [RD01]. We exported the DMP generated by means of DMP online as a doc file and we included it in this section. Section 2.1 contains it. Each part of the data management plan contains a set of questions (the one provided in the H2020 template) and the provided answers. The answers will be further elaborated in the next versions of the DMP when detailed information about the data used by I-REACT will be available.

2.1 DATA MANAGEMENT PLAN FOR THE I-REACT PROJECT (DMP V1.0)

ADMIN DETAILS

Project Name: I-REACT - Improving Resilience to Emergencies through Advanced Cyber Technologies

Grant Agreement: I-REACT is co-founded by the Horizon 2020 Framework Programme of the European Commission under grant agreement n. 700256

Principal Investigator/Researcher: Fabrizio Dominici

Project Data Contact: Fabrizio Dominici, dominici@ismb.it

Funder: European Commission (Horizon 2020)

1. DATA SUMMARY

Provide a summary of the data addressing the following issues:

- **State the purpose of the data collection/generation**
- **Explain the relation to the objectives of the project**
- **Specify the types and formats of data generated/collected**
- **Specify if existing data is being re-used (if any)**
- **Specify the origin of the data**
- **State the expected size of the data (if known)**
- **Outline the data utility: to whom will it be useful**

The I-REACT project aims at supporting emergency activities related to hazards (e.g., floods and fires) by integrating data from several external data sources. Specifically, the I-REACT project aims

at (i) supporting emergency activities related to hazards (e.g., floods and fires) and (ii) generating predictive and forecast models. To achieve the main goals of the project, several heterogeneous data must be properly collected, transformed and combined in order to tackle the problem from different perspectives. Specifically, publicly available open data and external data sources will be integrated with the content generated by the users of the I-REACT system (e.g., the user generated reports). The integration of several data sources will allow generating models based on complementary information that can be used to support the different activities related to hazards and emergency events. Both first responders and citizens will use and exploit, at different granularities, the data collected, transformed, and generated by I-REACT.

The collected data are mainly georeferenced data related to hazards and emergency events (floods and fires). Based on the initial analysis of the available data sources, and the use cases of the project, the following types of external data have been currently identified as potentially useful for the I-REACT project:

- Weather forecast data
- Radar and Satellite images/data
- Flood maps
- Copernicus land monitoring service data
- European flood awareness system data
- European forest fire information system data
- Disaster event historical data
- Global administrative area data
- Statistical data
- Social media data (e.g., tweets)

As reported in the above list, several external data sources will be re-used by I-REACT, each one providing a different facet of the hazard/emergency event we are facing. Both open and freely available data and non-public data will be used. Open and freely available data will be preferred.

In the I-REACT project, standard data formats (e.g., GeoJSON and Shapefiles), and metadata (e.g., INSPIRE complaint metadata) will be used in order to improve findability, accessibility, interoperability and re-usability of the data.

The I-REACT project will also generate data, based on the data (e.g., reports) generated by the users of I-REACT and the transformation and analysis of the collected data.

The main types of data generated by I-REACT are the followings:

- User generated reports
- Risk maps and weather forecast maps
- Flood and fire nowcasts and forecasts
- Climate change maps
- UAV imagery

Also, the data generated internally by I-REACT will be represented by means of standard data formats (e.g., GeoJson) and will be enriched with standard metadata.

The size of the data collected and generated depends on the considered data sources. The size of the datasets and files associated with the several data sources exploited by the project varies from some MBs to tens of GBs per file.

The collected and generated data will be useful for several end users and stakeholders. Specifically, the users of the I-REACT system (both citizens and first responders) will benefit from the data generated by the system. Moreover, third parties could also be interested in the data collected and generated by I-REACT, also for supporting decisions not directly related to the management of emergency events.

2. FAIR DATA

2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

We will share the public data generated by I-REACT in the Zenodo repository (https://www.zenodo.org/). Zenodo provides a set of basic functionalities that allows publishing data and searching them by means of keywords. Moreover, Zenodo automatically assigns a DOI to

each new uploaded dataset and allows specifying metadata, which can be profitably exploited to find the shared datasets.

The uploaded files will be identifiable and versioned by using a name convention consisting of project name, dataset name, version, and date.

For the data for which it is appropriate, the standard INSPIRE (http://inspire-geoportal.ec.europa.eu/) metadata will be used to enrich the shared data.

2.2 Making data openly accessible:

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

The sharing of the data will follow the ORD pilot principle "**as open as possible, as closed as necessary**". Specifically, the research data needed to validate the results presented in the published scientific papers will be made available. To support the sustainability of the project, according to the defined business plan, the data that are fundamental for the sustainability of project and are not used in the published scientific papers will not be disclosed.

The public datasets, and the related metadata, will be made available through the Zenodo repository. Standard data formats and metadata will be used to improve accessibility to the data by means of standard freely available tools.

2.3 Making data interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

We will use standard data formats (e.g., GeoJSON and Shapefiles) for representing the data used by I-REACT and standard metadata (INSPIRE complaint metadata) to improve interoperability and re-

usability of the data. This solution improves the interoperability of the modules of I-REACT and the interoperability with respect to external users interested in using the data of I-REACT.

2.4 Increase data re-use (through clarifying licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

The data collected from external sources, or generated transforming external data, will use the same license of the original data sources. For the subset of internally generated data that will be open, not based on external data sources, an open license will be used.

The research data associated with the published research papers will be made available as soon as the accepted papers will be published. The public research data published on Zenodo will be freely available as long as the Zenodo service will be available.

3. ALLOCATION OF RESOURCES

Explain the allocation of resources, addressing the following issues:

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long term preservation**

The public data will the published in the free Zenodo repository. Hence, there are no specific costs related to the storage of the public data.

Regarding the subset of non-publicly disclosed data exploited by the I-REACT project, the costs of the management of the data, and the related privacy issues, are already part of the costs related to the cloud architecture that will be used to implement the I-REACT system. Regarding the long-term preservation of the data, the costs for the management of the non-public data will be supported by the revenue associated with the exploitation of the data itself and the related services.

Project: I-REACT
Grant Agreement: 700256
Data Management Plan
Call ID: H2020-DRS-1-2015
Deliverable ID: D2.8
Page: 10 of 14

4. DATA SECURITY

Address data recovery as well as secure storage and transfer of sensitive data

Security and privacy are two important issues managed by the project. Specifically, Task "T2.4: Security & Privacy by design" addresses the security and privacy issues by adopting a security and privacy by design methodology. A detailed description of the selected methodology is reported in the deliverable "D2.4: Report on Privacy and Security". The solution described in D2.4 will be further specialized during the project based on the new privacy and security issues that will emerge during the project. Any further changes to the privacy and security solution that will arise during the project will be included in the following versions of the DMP.

Regarding the secure storage of the data, the project will use an architecture based on cloud services to store the data. The used services provide the functionalities needed to address secure storage and data security. Regarding sensible data, a secure and privacy by design methodology will be used to avoid the disclosure of sensible and personal data.

5. ETHICAL ASPECTS

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

According to the content of the ethics review and ethics section of DoA, the designed solution will imply the collection and the processing of personal data once operational. The management of such data will be regulated by a clear term of service, which must be read and accepted by users, and which will follow the EU data protection regulation. Note that data regarding financial details, sexual lifestyles, ethnicity, political opinion, religious or philosophical conviction will not be included in the Information Architecture. Different features of the system will rely on location-based technologies to determine the position of people and "objects" (such as infrastructures, resources, vehicles, etc.), which is crucial for the implementation of the I-REACT products. Regarding the localization of people, two different categories will be distinguished:

- **First responders** – who are the emergency operators and/or volunteers. Since they are professional users, they must accept the geolocation of their devices during the working time through a formal agreement with their employee. Their positioning is needed to send in-field reports and it is also aimed at monitoring their safety during the emergency response phase.
- **Citizens** – who must read and accept specific term of service in order to submit geolocalized reports through the I-REACT mobile application. Optionally, they can allow the I-REACT system

to locate their position, which will be shared only with subscribed authorities in case of emergency in order to effectively perform search and rescue operations. With respect to the data retrieved from Social Networks, the regulation included in their terms of service will be applied.

All personal data will be solely used for the project purpose, and they will be only accessible by the data owner. The geolocated reports will include only the user category and an ID that is not linkable with the user personal data. Personal data will not be subject of any exploitation and will not be distributed to any third party. Such data will be collected, stored, and processed following Privacy-by-design approaches in order to guarantee confidentiality, and anonymity. Thus, the data that will be publicly made available will be anonymized. I-REACT does not foresee the management of health data.

6. OTHER

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

Regarding the privacy issues, the data will be managed in compliance with the European regulation, as already described in Sections 4 and 5 of this Data Management Plan.

3 CONCLUSIONS

This deliverable describes how the data management plan of the I-REACT project has been generated and contains its first version. The DMP of I-REACT will be periodically updated, providing finer details, in time with the periodic evaluation/assessment of the project (i.e., at the mid-term review and at the final review).

Project: I-REACT
Grant Agreement: 700256
Data Management Plan
Call ID: H2020-DRS-1-2015
Deliverable ID: D2.8
Page: 13 of 14

END OF THE DOCUMENT

Project: I-REACT
Grant Agreement: 700256

Data Management Plan
Call ID: H2020-DRS-1-2015

Deliverable ID: D2.8
Page: 14 of 14