

European Tsunami Risk Service (unofficial demo website)
University of Naples Federico II, Italy



Coordinator
Fatemeh JALAYER¹
fatemeh.jalayer@unina.it

Data Management Plan

Version 20210723

DIGITAL ASSETS SUMMARY

Purpose of the provision of Data, Data products, Software, and Services

The European Tsunami Risk Service (henceforth called *the Service*) collects/generates Data, Data Products, Software, and Services (DDSS) to make them FAIR, i.e., Findable, Accessible, Interoperable, and Reusable. The main end-users for the DDSS are academic researchers, civil protection, Tsunami Service Providers (TSP), policymakers, and the private sector. Data and data products harmonization, interoperability, and provision are key aspects of the Service, which can be accessed at the following web address: <https://eurotsunamirisk.org/>

Types and formats of DDSS

The **data** collected are raw and/or processed data, e.g., observed damage versus tsunami flow depth, direct losses incurred due to tsunami, engineering demand parameters obtained from structural analysis. Data formats include text files containing comma separated values (.csv), hdf5, and .mat files. **Data products** (e.g., models, maps), such as fragility curves, vulnerability curves, consequence models, and recovery functions are created and/or collected by processing raw data, and/or exploiting existing models in the literature. The data and the data products collected/generated are the following: raw and/or processed data, tsunami fragility curves, consequence models, vulnerability curves, recovery functions, and tsunami risk maps. Data products for vulnerability are provided in .xml files in Natural hazard's Risk Markup Language, while the provided maps adhere to the Open Geospatial Consortium (OGC) standards. **Software** and **Services** are both made accessible and enable end users post-process raw data and generate data products (e.g., fragility curves). The **Services** are available online in Virtual Access through the users' web browser without requiring any additional software, while **Software** are provided as stand-alone open-source codes.

Information about re-usage of existing data

The Service will –to a high degree– be based on data currently residing in existing national and regional data and information repositories to produce new compilations and other derived data sets.

Origin of the data and the data products

The collected data and data products are originating from various sources, which include sensors, field surveys, public surveys, and publications.

¹ Associate Prof., Department of Structures for Engineering and Architecture (DiST), University of Naples Federico II.

Expected size of the data and the data products

It is not possible to estimate the size of the data, especially the size of the raw data. The size of the files for the data products will be relatively manageable. For example, the file size required to store the fragility curves for a type of structure amounts to less than 5 KB.

DSSS utility

The provided DDSS can be used by academic institutions, public agencies, policy makers, and the private sector.

1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

1.1 Discoverability of data and data products (metadata provision)

The webpage of the Service provides a searchable table listing the data and the data products available through the service.

1.2 Identifiability of DDSS

All DDSS provided in virtual access through the Service are identifiable by means of Digital Object Identifiers (DOI). The Service maintains a searchable table containing the provided DDSS and their corresponding DOIs. In the case of contributed data/data products, the DOI of the corresponding publication can be referenced on the table. The provided software is identified by the DOI for the software repository. The web services provided are also associated to the DOIs of the software repositories used to create the Docker images for the web services.

1.3 Reference naming conventions used

The naming convention for assets at risk follows the open exposure database (GED4ALL) schema for multi-hazard risk analysis.

1.4 Approach for clear versioning

The software and tools provided by the Service come with version numbers. The data and data products do not come with version numbers, since they are considered static and tied to a DOI. In the case of correction of data or data products new DOIs will be used.

1.5 Standards for metadata creation

The standards for metadata creation used follow the OCG standards and the EPOS-DCAT-AP metadata template.

2 MAKING DATA AND DATA PRODUCTS OPENLY ACCESSIBLE

2.1 Data and data products that are made openly available

All provided DSS are made accessible for free and openly. The only prerequisite is a free user registration, which has the sole purpose of evidencing the utility of the Service. Future versions of the Service may consider keeping some data and data products closed or purging data and data products for reasons such as embargo periods before publication, and national safety. In that case, the Data Management Plan (DMP) will be updated.

2.2 Access to DDSS

The Service provides Virtual Access to DDSS. The users can access online the data, the data products, and the software. Moreover, the users can download data/data products/software for offline use. In specific, the Service links to a repository on

zenodo.org, where data/data products are stored. The service also links to repositories on github.com, where the software made available by the service is located. As far as the services are concerned, they have the form of online Docker images hosted at mybinder.org.

2.3 Methods, software, and tools required to access the data and the data products

Any web browser can be used to view and download the files of the data and the data products. To access the information contained in the files, software that can read plain text files, text files containing comma separated values, xml files in Natural hazard's Risk Markup Language, hdf5 files, or .mat files.

2.4 Location of the repositories for the data and the data products

The data and the data products are located in repositories on zenodo.org, while the software is located in repositories github.com. Copies of the software on github.com are located in repositories on zenodo.org. The website of the Service provides links to the repositories where the data/data products/software are located.

2.5 Access in case of restrictions

Currently, there is no restriction of access to DDSS. In the case that such restrictions are put in place, the DMP will be updated, and access will be given through a page of the website of the Service accessible only to registered users. Registered users will be granted access to data/ data products they request after presenting evidence that they meet the criteria for access to the restricted data/data products.

3 MAKING DATA AND DATA PRODUCTS INTEROPERABLE

3.1 Data and data products interoperability

The Service has a very strong focus on interoperability of the data and the data products it makes available. Data products, such as the fragility and vulnerability curves provided by the Service come in xml files in the Natural hazard's Risk Markup Language following the GED4ALL schema. The DMP will be updated as more data and data products are made available by the Service. For spatial data, the use of the Open Geospatial Consortium standards is foreseen. Maps are provided following the OpenGIS Web Map Tile Service Implementation Standard or as Web Map Services. Raw data is provided in interoperable file types, i.e., .csv and .mat files. Software interoperability is emphasised, and the provided software is open source and in widely used programming languages (e.g., Python, MATLAB). The provided web services require only a web browser to be accessed and require/generate data in the interoperable data formats adopted by the Service.

3.2 Use of standard vocabularies for all datatypes

European and international standards for the storage, exchange, and dissemination of project data. INSPIRE (the European Directive on Infrastructure for Spatial Information) compliance will be used wherever possible. Where this is not possible, Commission for the Management and Application of Geoscience Information (CGI) standards will be used. Standard vocabularies will be used to the extent that they exist or will be developed.

4 INCREASING DATA AND DATA PRODUCT RE-USE THROUGH CLARIFICATION OF LICENSES

4.1 Data/ data products licenses for widest reuse possible

The Service releases data and data products with a Creative Commons Attribution 4.0 International (CC BY 4.0) license. This is a “Free Culture” license, it allows commercial use, and it does not impose a restriction on the type of license of the derivative works. These conditions offer the maximum possibility of reuse while still requiring attribution of credit to the creators of the parent work. The details of this license are found on the following web page:

<https://creativecommons.org/licenses/by/4.0/>

4.2 Time of data/data products release for re-use

The data will be made available immediately. The possibility of release after the expiry of an embargo period may be added in future versions of the Service.

4.3 Reuse of data/data products by third parties

The data/ data products provided by the Service are reusable by third parties according to the conditions of the license.

4.4 Data quality assurance processes

The service checks the provided data products for adherence to current standards. Further data quality assurance processes will be described in future versions of the DMP.

4.5 The length of time for which the data remain re-usable

Data and data products remain reusable for an undetermined amount of time, as this is not covered by the license.

5 ALLOCATION OF RESOURCES

5.1 Costs of making data FAIR

In accordance with the EPOS Data Policy, the Service makes DDSS available with emphasis on their Findability, Accessibility, Interoperability, and Reuse (FAIR - <https://www.go-fair.org/>). The Service is unable to provide an estimation of these costs in this version of the DMP.

5.2 Describe how you intend to cover these costs

The costs are covered by the University of Naples Federico II.

5.3 Data Managers

Konstantinos TREVLOPOULOS² is currently serving as the Data Manager of the Service.

5.4 Describe costs and potential value of long-term preservation

Currently the costs of long-term preservation are unknown. The potential value is great, as the Service is an infrastructure that enables sharing and reusing data and data products.

² University of Naples Federico II, Department of Structures for Engineering and Architecture Via Claudio 21, Naples 80125, Italy. konstantinos.trevlopoulos@unina.it

6 DATA SECURITY

6.1 Safe long-term storage security of data/data products/software

Data/data products/software are stored on zenodo.org “in CERN’s Data Centre for as long as CERN exists”.

6.2 Digital Continuity Plan/Disaster Recovery Plan

A backup copy of the data/data products/software is kept at the University of Naples Federico II (Department of Structures for Engineering and Architecture, Via Claudio 21, Naples 80125, Italy).

7 ETHICAL ASPECTS

The Service requests, processes and stores no personal data. The Service is not liable for any data that may be processed or stored by the providers of the web services used by the Service.

8 OTHER ISSUES

Future versions of the DMP will state if and which other national/funder/sectorial/departmental procedures are used for data management.