Plan Overview

A Data Management Plan created using DMPonline

Title: Database of Individual Seismogenic Sources

Creator:Roberto Basili

Principal Investigator: Roberto Basili

Data Manager: Roberto Basili

Affiliation: Other

Template: DCC Template

ORCID iD: 0000-0002-1213-0828

Project abstract:

DISS is a georeferenced repository of tectonic, fault, and paleoseismological information expressly devoted, but not limited, to potential applications in analyzing earthquake hazards at the regional and national scale. It represents faults in 3D; all its records are fully parameterized; it tends to completeness.

The core objects of DISS are as follows:

- The individual seismogenic source is a simplified three-dimensional representation of a rectangular fault plane. Individual seismogenic sources are assumed to exhibit "characteristic" behavior concerning rupture length/width and expected magnitude.
- The composite seismogenic source is a simplified three-dimensional representation of a crustal fault containing an unspecified number of seismogenic sources that cannot be singled out. Composite seismogenic sources are not associated with specific earthquakes or earthquake distributions.
- The debated seismogenic source is an active fault proposed in the literature as a potential seismogenic source but was not considered sufficiently constrained to be included in any of the other categories.
- The subduction source, a simplified and three-dimensional representation of the complex subduction system, is mainly identified by the depth contours of the subducted slab. Similarly to composite seismogenic sources, subduction sources are not associated with specific earthquakes or earthquake distributions.

ID: 110060

Start date: 01-01-1999

Last modified: 16-11-2022

Grant number / URL: Joint Research Unit EPOS-Italy, programme of activities 2021-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any

relationship to, your project or proposal

Database of Individual Seismogenic Sources

Data Collection

What data will you collect or create?

The compilation of DISS will exploit the wealth of information available from Earth Science studies, including, but not limited to, earthquake geology, seismology, seismotectonics, and geodynamics. The main purpose of such datasets is the geometric reconstruction of potential earthquake sources and estimating their activity rates. Re-used data will mainly come from the scientific literature.

The generated dataset will be an integrated data product from complex analyses or community-shared data harmonization.

The dataset will be compiled and distributed using open-source GIS software and open file formats. The distributed data volume will be limited to about one gigabyte, which will pose no problem for long-term preservation and access.

How will the data be collected or created?

Generated and re-used data will most often be geospatial data providing the location of potential seismogenic faults, their geometry, and their behavior. Parameters detailing geometry and behavior will be linked to the spatial data in tabulated attributes. There is no community standard for this type of data; however, the dataset structure will consider prescriptions dictated by the needs of the earthquake hazard and seismotectonic modelers.

For data available for download, there will be a different folder for each format. Each folder will be named with the dataset acronym and version followed by the format's name (e.g., DISS_3.3.0_tab). Each file will be named with the specific dataset acronym (e.g., ISS or CSS), an abbreviation identifying the subset and version, and the proper format extension (e.g., ".tab"). Details about the content of each specific dataset will be provided in a README file added to each file folder. The same naming rules will be applied to the OGC web service layers.

The quality control of the distributed data will be carried out according to a multi-step workflow described in the data quality assurance document available in the documentation section of the EDSF portal.

Documentation and Metadata

What documentation and metadata will accompany the data?

All DISS versions will be accompanied by comprehensive documentation addressing the data structure, the definition of variables, and the units of measurement.

Metadata will be openly available and contain enough information (direct link) to enable the user to access the data.

Provisions for metadata will include:

- metadata offered with the DOI as required by DataCite;
- metadata offered through the INGV Open Data Portal;
- metadata offered through the standard Open Geospatial Consortium (OGC) protocol CSW.
- INSPIRE for the versions mapped in the Italian "Repertorio Nazionale dei Dati Territoriali."

Ethics and Legal Compliance

How will you manage any ethical issues?

There is no ethical reason that could impact data distribution and sharing. A disclaimer will be associated with the dataset to

remove legal liability from the data owner and publisher. Users will also be cautioned to carefully consider the dataset's nature before using it for decisions concerning personal or public safety or business involving substantial financial or operational consequences.
No personal data will be collected or distributed with the dataset.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

All DISS versions will be distributed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license terms. Users can request additional permissions to use the dataset by contacting the DISS Working Group.

Storage and Backup

How will the data be stored and backed up during the research?

The data will be stored in the server that publishes the static file and in the server that issues the OGC WMS and WFS standards.
The data will be backed up using a storage server connected to the INGV private network.
To back up the database, we will use the standard PostgreSQL tool "pg_dump."
Since DISS versions will be released at irregular intervals in the order of months to years, there is no need to schedule an automatic backup procedure.
The entire website where DISS is published is regularly backed-up.
The responsible for the backup and recovery procedure is Roberto Vallone (INGV).
In case of an incident with the publishing server, data will be recovered by restoring the database and the files from one of the multiple backup services. In particular, the database will be restored using the standard "pg_restore" tool of PostgreSQL.

How will you manage access and security?

All DISS versions will be openly accessible.
SSL transfer for HTTP (HTTPS) is implemented and is chosen per default for all hosted services on the EDSF Installation where DISS is published.
No sensitive data will be stored.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

DISS data and metadata stored in the INGV repositories will remain available indefinitely.
DISS is an integrated data product; as such, most of the raw and processed data used to compile the DISS datasets will remain with their owners. Whenever possible, part of the data will accompany the documentation of individual database records and be made available through the database user interface.
DISS will be used to devise the input datasets to carry out earthquake hazard analyses (e.g., ground shaking or tsunami) and earthquake scenarios or to develop seismotectonic and geodynamic models.

What is the long-term preservation plan for the dataset?

Datasets will be deposited in two servers owned by INGV, installed in two different institutional premises for security reasons. Since each DISS version will occupy less than 100 MB of disk storage and the file formats used will presumably be of common use for many years, the cost of storage related can be considered easily affordable.

Data Sharing

How will you share the data?

The standard OGC protocols WMS and WFS will be adopted to guarantee interoperability with other datasets or spatial data.
The availability of the DISS versions as downloadable files in popular formats (ESRI shapefiles, MapInfo Tables) will facilitate users combining and analyzing DISS with other geographically referenced data in a desktop Geographic Information System (GIS).
A different DOI will identify each DISS version.

Are any restrictions on data sharing required?

The entire collection of DISS versions will be made openly accessible with no restrictions except for properly using the citation prescribed by the attribution license.

Responsibilities and Resources

Who will be responsible for data management?

The persons responsible for the data management, curation, preservation, and distribution are the members of the DISS Working Group as indicated on the DISS web portal.

What resources will you require to deliver your plan?

Storage, archiving, re-use, and security costs will partly be covered by EPOS and INGV institutional funding. When additional resources are necessary, they will be sought through project funding.