
Manuale operativo del portale dati Open Access OE - DataRep

Release 1.0

INGV Osservatorio Etneo - UFSOIT

13 dic 2022

1	Introduzione	3
1.1	Come pubblicare sul portale DataRep	3
2	Processo di caricamento dati sul portale DataRep	5
2.1	Caricamento del <i>dataset</i> sullo spazio Google Drive personale	6
2.2	Creazione della <i>scheda descrittiva</i> dal Metadata Editor	7
2.3	Compilazione dei <i>Metadati</i>	8
2.4	Selezione dell'elenco di file da pubblicare	10
2.5	Validazione della scheda e assegnazione del DOI (Ufficio Gestione Dati)	13
2.6	Pubblicazione sul portale DataRep	14



DataRep

INGV - OSSERVATORIO ETNEO
OPENDATA Repository

Questo documento rappresenta un manuale operativo di utilizzo del Portale Dati Open Access **DataRep**.

Indice:

Il repository **DataRep** fornisce uno strumento che permette, al personale del *Osservatorio Etneo* dell’Istituto Nazionale di Geofisica e Vulcanologia, di pubblicare datasets rispettando le politiche del paradigma **Open Access**.

Questo strumento, si occupa di creare automaticamente le pagine web: *landing pages* che ospiteranno sia il dataset che i metadati ad esso associati, rendendoli pubblicamente fruibili.

A ciascun dataset è associato un identificativo univoco detto **DOI**: Digital Object Identifier, la *landing page*, creata dal portale alla pubblicazione di un dataset, garantirà la risolvibilità del DOI.

1.1 Come pubblicare sul portale DataRep

Il processo di pubblicazione di un dataset su **DataRep** coinvolge due entità:

1. il **Metadata Editor** ([link](#)) collegato al **Registro Dati** istituzionale e gestito dall’**Ufficio Gestione Dati** dell’**INGV** ([link](#));
2. il portale **DataRep** ([link](#)) dove i dataset vengono resi disponibili nel rispetto delle **FAIR** policies dettate dal paradigma Open Access.

Entrando nel dettaglio delle entità coinvolte nel processo di pubblicazione, il **Metadata Editor**, si occupa:

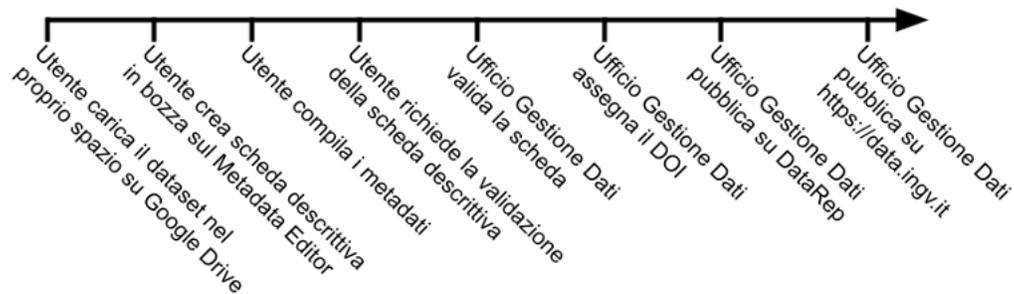
- della gestione centralizzata dei metadati che descrivono i datasets
- pubblica automaticamente la scheda del dataset sul portale **DataRep**
- carica i dati sul portale DataRep.

Il portale **DataRep** è un servizio web che offre alla comunità scientifica un mezzo per condividere e conservare i risultati scientifici e che ne consente la *riproducibilità* ed il *riutilizzo*:

- associa automaticamente i metadati della pubblicazione generati dal Metadata Editor alle landing pages;
- implementa la risolvibilità del DOI associato alla pubblicazione;
- permette l’esportazione secondo vari schemi di metadati tra cui quello, definito da **DataCite** ([link](#));

Processo di caricamento dati sul portale DataRep

Il processo da seguire per pubblicare *dataset* sul repository **DataRep** coinvolge due entità: il *Metadata Editor* ed il *DataRep*, ed è descritto per linee generali nella seguente figura e nei seguenti punti:



Utente

- carica il *dataset* nel proprio spazio su **Google Drive**
- crea la **scheda descrittiva**
 - compila i *metadati*
 - seleziona i *file* da pubblicare
- richiede la validazione della scheda

Ufficio Gestione Dati (UGD)

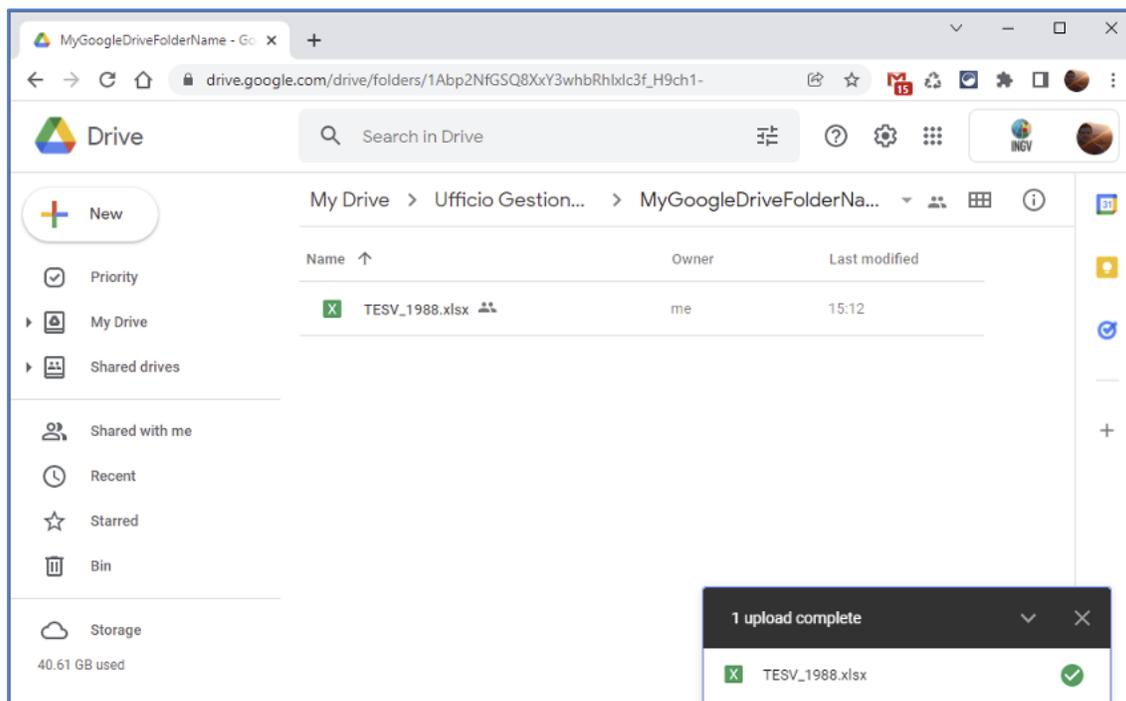
- procede alla validazione della scheda
- verifica se assegnare il *DOI*
- pubblica su **DataRep**

2.1 Caricamento del *dataset* sullo spazio Google Drive personale

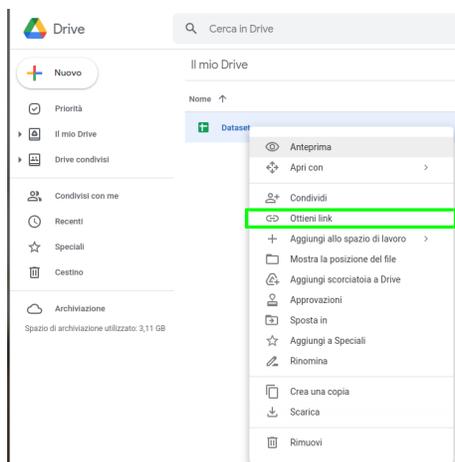
Inizialmente l'utente carica il *dataset* all'interno di uno spazio dati personale e privato sul cloud di **Google Drive**.

I passi necessari sono illustrati nelle seguenti figure.

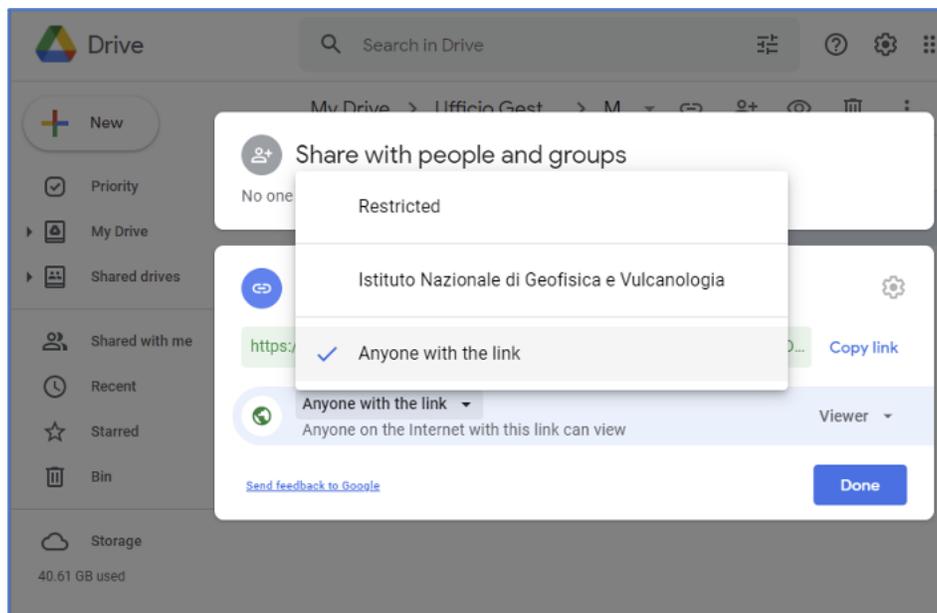
Tramite il proprio spazio personale istituzionale sul Google Drive, l'utente carica il dataset da pubblicare sul portale DataRep.



Successivamente l'utente rende accessibili gli stessi files del dataset mediante la funzionalità «*Otteni link*» o «*Get link*».



È necessario selezionare l'opzione «*Chiunque abbia il link*» o «*Anyone with the link*» generando univocamente un link per ciascun file che sarà utilizzato dal Metadata Manager per l'operazione di upload su DataRep.



Un dataset può essere composto da uno o più file a cui corrisponderanno di conseguenza più *link*.

È consigliato archiviare in un unico file di tipo **zip** o **tar, rar** i differenti *file*, così da creare un unico link di condivisione.

2.2 Creazione della *scheda descrittiva* dal Metadata Editor

Ottenuto il link di condivisione, l'utente si autentica sul portale del Metadata Editor utilizzando il proprio account istituzionale (*nome.cognome@ingv.it* come per l'accesso alla posta elettronica) al fine di **creare** una scheda descrittiva del *dataset*:

1. Il primo passo consiste nella *creazione* di una nuova scheda descrittiva in **bozza - Drafts**, scegliendo se iniziare da una *scheda vuota* o dalla *copia* di una scheda già esistente da completare;
2. A questo punto il sistema consente di inizializzare una nuova bozza, permettendo all'utente di scegliere se creare un nuovo *record* che descrive il *dataset*, a cui: **a)** non è ancora stato assegnato nessun DOI, **b)** è già stato assegnato un DOI.

Create a new record describing data that IS NOT associated with a DOI. INGV will assign a DOI using this editor.

Compile a new record starting with empty fields

Alternatively, you may enter a new record starting from a set of pre-compiled metadata, you have two choices

Copy and then modify metadata from another Record ID

or

Copy and then modify metadata from another DOI

Create a new record describing data that IS ALREADY associated with a DOI that was not assigned by INGV.

Existing DOI

Please, note that most of the metadata associated to DOI not assigned by INGV cannot be changed.

Per dataset a cui bisogna assegnare un nuovo DOI

Per dataset a cui è già stato assegnato un DOI da altra organizzazione

3. Il *web-form* verrà mostrato all'utente come un'anteprima e sarà compilabile solo in parte in quanto la scheda vera e propria non verrà creata nel sistema fino a quanto l'utente non avrà cliccato il bottone «Create» in fondo alla scheda.

2.3 Compilazione dei *Metadati*

Scelta la modalità di creazione del nuovo record (paragrafo precedente), il sistema mostrerà un *web-form* con diversi campi da compilare che costituiscono i *metadati* caratteristici di una pubblicazione scientifica.

Solo dopo aver cliccato il bottone «Create» in fondo alla pagina la scheda sarà creata nel sistema e sarà editabile in ogni sua parte.

Quando la scheda viene creata sarà inserita nel metadata editor sotto forma di bozza (*draft*) e sarà considerata tale fino alla fine del processo di validazione, processo che verrà avviato dall'utente quando avrà terminato la compilazione della scheda.

L'utente descriverà il dataset compilando esaurientemente quanti più campi possibili tra quelli disponibili come nella scheda di esempio della seguente figura.

Metadata contained in **blue forms** must be compiled.
Click here [to show the list of required fields.](#)

▼ **Basic information**

Warning: **Title** and the **Year of publication** cannot be changed in any way (not even the order!) once this record will be validated and it will be moved into the Data Registry as these fields will be used for generating the bibliographic citation.
The requirements for the landing page are listed in the [INGV Data Policy](#).

Title - Include the acronym -if any- in the title to make it appear in the bibliographical citation

Title language

Acronym Version

Year of publication - Format YYYY. In case of dynamic data, enter the first year in which the data were published online

Landing page [to show the list of required fields.](#)

Status Data Census ID 2016-2017 [to show the list of required fields.](#)

Data group - click for more info [to show the list of required fields.](#)

► **Other titles**

▼ **Data producers**

Click here [to show the list of required fields.](#) to read the definition of the available roles that can be associated to each contributor.

Main authors must be always compiled.
A person (or organisation) cited as a main author, might be additionally cited among the contributors too, if you want to specify its role.
Mind that only the list of main authors will be used to compile the automatically generated bibliographical data citation.
If you are unable to find a person or an organisations among the available ones, add the name, surname, and affiliation in the "General notes" below and send us an email to ufficioquestionedati@inav.it

Nota: In particolare, il campo "Landing page" non deve essere compilato, perché questa informazione verrà restituita automaticamente dalla logica del portale DataRep non appena, verrà fatta richiesta di pubblicazione.

Nella compilazione della scheda, dovranno essere indicati anche un *referente scientifico* e uno *tecnologico* (le due figure possono coincidere). I referenti andranno indicati nella sezione dei metadati denominata “Info specific to INGV”. Oltre ai metadati obbligatori, evidenziati in blu nella scheda (Fig. sopra), si suggerisce di compilare anche informazioni utili come la copertura temporale, geografica e soprattutto eventuali relazioni con altri dati o pubblicazioni nell'apposita sezione “Relations to other data and publications”.

Infine, è anche possibile caricare un'immagine di copertina in fondo al *web-form* alla scritta “upload a cover image file”.



Durante la compilazione dei campi è possibile verificarne la correttezza in ogni momento agendo sul bottone “Check if this record can be validated”, posizionato in alto sulla form di caricamento dei metadati.



Se il processo rileva errori o campi obbligatori non compilati, tutte le anomalie saranno dettagliatamente elencate dal *Metadata Editor* e l'utente potrà procedere alla loro risoluzione.

Infrasonic signals recorded at Mt. Etna stations during period 15-19 January 2021 (ISGN) [ID 687] Print

Result of the automatic check of metadata: **this record is not valid.**

Basic information

Warnings:

- Data Census ID 2016-2017:** This record seems to be related to data that did not appear during the 2016-2017 data survey, so it must be extensively validated and approved by the involved Section's Directors, and then by the involved Department's Directors.

Data Access and Distribution

Errors:

- Target repository where data will be uploaded:** At least one file should be provided.

Relations to other data and publications

Warnings:

- Relations to other data and publications:** There is no relation to other data or publications.

Cover image

Warnings:

- Cover image:** The cover image is missing.

Si consiglia di salvare frequentemente le modifiche durante la compilazione cliccando sul bottone "Save" in fondo alla scheda per evitare il rischio di perdere informazioni già inserite. La scheda può essere compilata a più riprese sia da parte dal referente scientifico che dal referente tecnologico, ma non nello stesso istante.

Se è necessario permettere ad altre persone di editare la scheda, queste possono essere aggiunte dai referenti nella sezione “Additional users allowed to update

metadata of this record” nella sezione di metadati “Info specific to INGV”.

Quando la compilazione della bozza della scheda sarà ultimata e il sistema di verifica automatico non rileverà più errori, l'utente potrà avviare il processo di validazione cliccando sul bottone "Start the validation process" all'inizio della scheda.

2.4 Selezione dell'elenco di file da pubblicare

A questo punto l'utente aggiunge l'elenco dei file da caricare su DataRep:

Il Metadata Editor mostrerà le funzioni di caricamento sul portale DataRep quando le due condizioni seguenti saranno soddisfatte:

1. Il **dataset** **deve** essere dichiarato **statico** nella sezione di metadati "Details about data". Se necessario, in futuro sarà possibile pubblicare nuove revisioni, pubblicando un nuovo dataset seguendo la stessa procedura con l'accortezza di indicare la relazione con la versione precedente nella sezione di metadati "**Relations to other data and publications**". Nell'immagine in basso un esempio di selezione del tipo in statico.

Details about data

Type of element:
- Data collection (or Database). An aggregation of resources, which may encompass collections of one resourceType as well as those of mixed types. A collection is described as a group; its parts may also be separately described. Raccolta di dati (o "Set di dati") indipendenti sistematicamente o metodicamente disposti e individualmente accessibili mediante mezzi elettronici o in altro modo. La tutela delle banche di dati non si estende al loro contenuto e lascia impregiudicati diritti esistenti su tale contenuto. Dalla definizione è esclusa l'infrastruttura informatica che gestisce la banca dati, ivi compresi eventuali Servizi per l'interazione con i dati. [[more info](#)]
- Dataset. Data encoded in a defined structure [[more info](#)].

Data Management Plan (DMP). A DMP describes the data management life cycle for the data to be collected, processed and/or generated and it should include information on: 1) the handling of research data during & after the end of the project, 2) what data will be collected, processed and/or generated, 3) which methodology & standards will be applied, 4) whether data will be shared/made open access and, 5) how data will be curated & preserved (including after the end of the project). You may find a useful tool for setting up your DMP at the [Digital Curation Centre \(DCC\)](#).

Dati di monitoraggio. Dati provenienti da sistemi osservativi istituzionali, acquisiti o prodotti con ricorrenza e sistematicità per la caratterizzazione e la comprensione dei processi fisici e/o chimici del sistema Terra.

Dati di sorveglianza. Dati elaborati per scopi di gestione dei rischi naturali, di protezione civile o di pubblica utilità forniti a enti e/o istituzioni pubbliche o private, nazionali e/o internazionali, a seguito di accordi formali.

Sizes: for example bytes, pages, inches, or anything related to the duration (extent), like hours, minutes, days, etc., of the data.

Data Level
--

Data Class
--

Data theme - See [EUROVOC](#)
none

This list can only be changed after you have created the record (i.e. after you have pressed "CREATE")

Static or dynamic data
 --
 static, data will never change
 dynamic, data will be updated

Type of updates - Only compile when the data is dynamic, do not compile when static:
--

Nota: Con statico si intende che il dataset non verrà più aggiornato dopo la sua pubblicazione.

2. Almeno una persona elencata tra i "**Main authors**" della sezione di metadati "Data producers" deve essere **affiliata all'Osservatorio Etneo**.

Data producers

[Click here](#) to read the definition of the available roles that can be associated to each contributor.

Main authors must be always compiled.
A person (or organisation) cited as a main author, might be additionally cited among the contributors too, if you want to specify its role.
Mind that only the list of main authors will be used to compile the automatically generated bibliographical data citation.
If you are unable to find a person or an organisations among the available ones, add the name, surname, and affiliation in the "General notes" below and send us an email to ufficiogestionedati@ingv.it

Warning: the list of Main authors cannot be changed once this record will be validated and it will be moved into the Data Registry as this list will be used for generating the bibliographic citation.

Main authors
none

Contributors
none

This list can only be changed after you have created the record (i.e. after you have pressed "CREATE")

This list can only be changed after you have created the record (i.e. after you have pressed "CREATE")

Le funzioni per caricare dati su DataRep saranno mostrate solo dopo che l'elenco degli autori sarà compilato.

E' possibile verificare in ogni momento le affiliazioni di ciascun autore cliccando sul bottone *Check affiliations* sotto l'elenco degli autori.

Click to check all affiliations of both main authors and contributors

[Check affiliations](#)

Taking into account main authors and contributors, in 2022 these are the organisations involved:

Organisation	People
Istituto Nazionale di Geofisica e Vulcanologia (INGV) ↗	Cassisi, Carmelo (begin 2013-12-01, end --) Sciotto, Mariangela (begin 2019, end --) Scuderi, Luciano (begin 2001, end --) Torrìsi, Orazio (begin 2001, end --)

Taking into account the list of main authors, in 2022 these are the INGV Sections involved:

INGV Section	Director	People affiliated
Istituto Nazionale di Geofisica e Vulcanologia (INGV) - Osservatorio Etneo (OE)	Branca, Stefano	Cassisi, Carmelo Sciotto, Mariangela Scuderi, Luciano Torrìsi, Orazio

In addition to the involved Section(s) Director(s) above, and taking into account the INGV Departments declared, these are the other people involved in the validation process of this Data Registry record:

Department	Reference person in the Data Management Office	Department Director
Vulcani	Montalto, Placido Maria	Neri, Augusto

Ciascun autore affiliato all’INGV può aggiornare direttamente le proprie informazioni (affiliazione compresa) modificando il proprio profilo nella sezione “People” del Metadata Editor. Le modifiche apportate ai profili personali saranno monitorate.

A questo punto è possibile procedere alla selezione dei *file* da pubblicare su **DataRep** che erano precedentemente stati caricati sul **Google Drive**. Per fare ciò:

1. espandere la sezione “Data Access and Distribution” di metadati;
2. selezionare *Osservatorio Etneo Open Data Repository* nel campo “Target Repository where data will be uploaded”;

Data Repository

Conditions: Data must be static and at least one main author must be affiliated to Osservatorio Etneo.

Target repository where data will be uploaded

Osservatorio Etneo Open Data Repository ▾

Files to be uploaded in the Data Repository via Google Drive

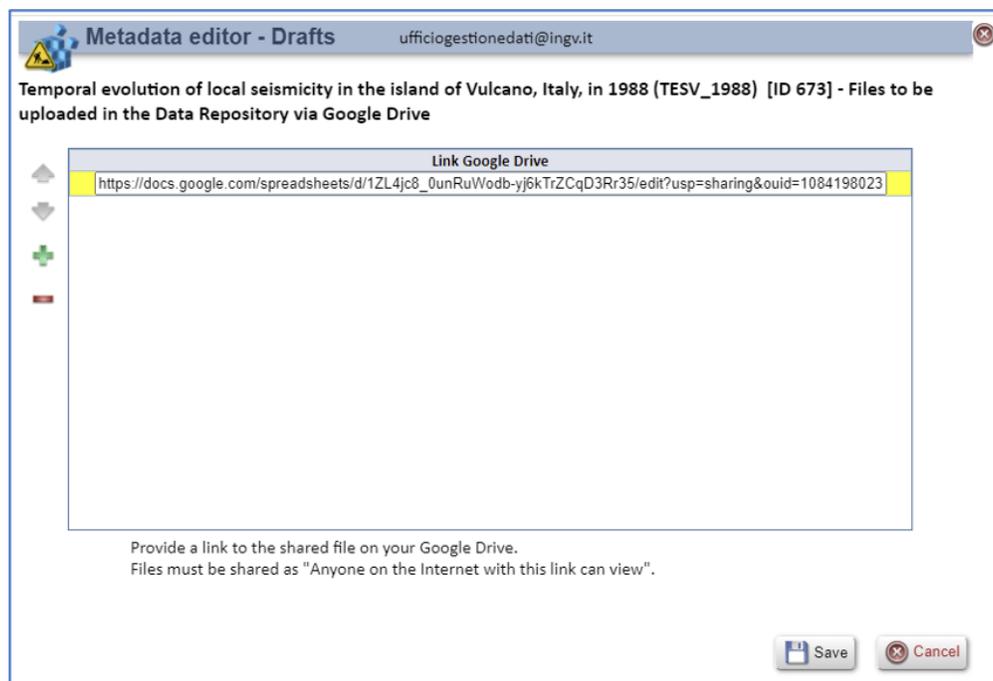
none

[Modify the list](#)

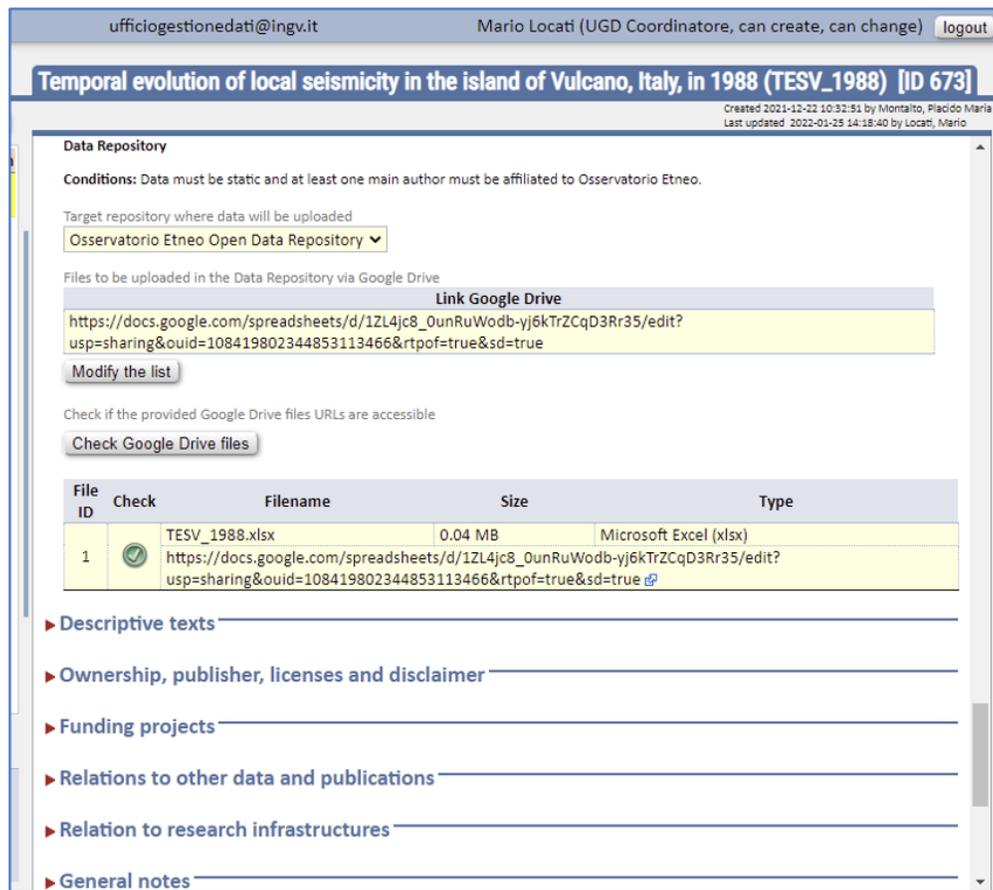
Check if the provided Google Drive files URLs are accessible

[Check Google Drive files](#)

3. cliccare il bottone “Modify the list”, per aprire la maschera di selezione dei files da Google Drive;
4. inserire i link, generati come visto nel paragrafo like *Caricamento del dataset sullo spazio Google Drive personale*, a ciascun file



5. per verificare che i dati siano caricabili su DataRep cliccare sul bottone “Check Google Drive files”, la spunta verde in corrispondenza della colonna *Check* indica che sarà possibile procedere al caricamento;



6. infine cliccare sul bottone *SAVE* in fondo alla scheda per salvare la Bozza creata.

2.5 Validazione della scheda e assegnazione del DOI (Ufficio Gestione Dati)

Quando la compilazione della scheda sarà completata si potrà procedere alla **validazione** cliccando sul bottone “Start the validation process”.

Avvertimento: in caso di problemi, l’Ufficio Gestione Dati si occuperà di contattare l’utente per eventuali correzioni o integrazioni.

La figura seguente mostra gli stati di avanzamento del processo di validazione.

The screenshot shows the validation progress for a record titled "AM(eri)Go: alla scoperta degli effetti dei terremoti nelle Marche [ID 425]". The interface is divided into three main sections for approval:

- 1. Data Management Office:**
 - Coordinator: Locati, Mario (2020-10-22 17:21:07)
 - Terremoti: Aste, Francesco Mariano (2020-10-23 09:53:37)
- 2. Section Director:**
 - Osservatorio Nazionale Terremoti (ONT): Stramondo, Salvatore (2020-10-26 17:14:45)
 - Sezione di Bologna: Costa, Antonio (2020-10-26 17:06:17)
 - Sezione di Milano: Lodi, Lucia (2020-10-23 12:34:19)
- 3. Department Director:**
 - Terremoti: Di Giovambattista, Rita

Below these sections, there is a checkbox for "Check if this record can be validated" and a "TOP PRIORITY" section. The "Basic information" section includes a warning about title and year changes, and input fields for Title, Title language (Italian), Acronym (AM(eri)Go), Version (1.0), and Year of publication (2019).

La validazione potrebbe arrivare fino a tre gradi di approvazione a seconda delle caratteristiche del dataset e degli autori indicati:

- il primo grado di validazione è in capo all’*Ufficio Gestione Dati*, nello specifico il coordinatore e un referente per ciascuno dei Dipartimenti indicati;
- il secondo grado è in capo ai *Direttori delle Sezioni* coinvolte identificate automaticamente dal Metadata Editor in base alle affiliazioni disponibili nella sezione “People” corrispondenti agli autori elencati;
- il terzo grado di validazione è in capo ai *Direttori dei Dipartimenti* indicati in “INGV Department” nella sezione di metadati “Info specific to INGV”.

Nota: se il dataset è una nuova versione di un dataset già validato e pubblicato in precedenza, si attiverà automaticamente solo il primo grado di validazione.

2.6 Pubblicazione sul portale DataRep

L'ultimo step, nel processo di upload di un dataset sul portale DataRep, è demandato ai responsabili dell'Ufficio Gestione Dati. L'interfaccia web del *Metadata Editor*, per queste figure, mette a disposizione due bottoni, mostrati nella figura sottostante.

- Il bottone "Upload a draft" caricherà su DataRep una bozza della scheda, contenente dati e metadati senza rendere la *landing-page* pubblicamente accessibile.
- Il bottone "Upload and publish or edit an already published record", caricherà su DataRep la scheda contenente dati e metadati e pubblicherà automaticamente la *landing-page*.

Data Repository

Experimental: At the moment this feature is available if at least one main author is affiliated to Osservatorio Etneo.

Target repository where data will be uploaded
Osservatorio Etno Open Data Repository

Date of first publication in the repository Date of the last update sent to the repository

Files in the Data Repository

Link Google Drive	File name	Size in MB	Type	Upload date
<a data-bbox="224 913 1031 934" href="https://docs.google.com/spreadsheets/d/1ZL4jc8_0unRuWodb-yj6kTrZCqD3Rr35/edit?usp=sharing&oid=10841980234485311346[...]">https://docs.google.com/spreadsheets/d/1ZL4jc8_0unRuWodb-yj6kTrZCqD3Rr35/edit?usp=sharing&oid=10841980234485311346[...]	TESV_1988.xlsx	0.04	Microsoft Excel	--

Modify the list

Check if the provided Google Drive files URLs are accessible
Check Google Drive files

File ID	Check	Filename	Size	Type
1		TESV_1988.xlsx <a data-bbox="316 1081 1128 1102" href="https://docs.google.com/spreadsheets/d/1ZL4jc8_0unRuWodb-yj6kTrZCqD3Rr35/edit?usp=sharing&oid=10841980234485311346[...]">https://docs.google.com/spreadsheets/d/1ZL4jc8_0unRuWodb-yj6kTrZCqD3Rr35/edit?usp=sharing&oid=10841980234485311346[...]	0.04 MB	Microsoft Excel (xlsx)

Create the record and upload files only Create the record, upload files and publish

Upload a draft Upload and publish or update an already published record

La *landing-page* creata dal Portale DataRep è mostrata nella figura sotto

The screenshot displays the DataRep portal interface for a dataset. At the top, the INGV logo and 'OPENDATA Repository' are visible. The dataset title is 'Temporal evolution of local seismicity in the island of Vulcano, Italy, in 1988 (TESV_1988)'. The author is listed as Falsaperla, Susanna. The DOI is 10.13127/vulcano/tesv_1988. The dataset is categorized as 'earthquakes' and 'seismic events related to fluid dynamic'. A file named 'TESV_1988.xlsx' is available for download, with a size of 36.1 kB. The page also shows a 'Share' section with social media icons and a 'Cite as' section with a citation style input field.

In particolare è possibile notare come la pagina creata conterrà:

- un estratto dei metadati inseriti sul Metadata Editor,
- il DOI assegnato alla pubblicazione,
- il dataset,
- la possibilità di esportare i metadati nei diversi formati supportati dal portale.

Una email notifica gli amministratori del portale che una nuova pubblicazione è stata effettuata ed è disponibile; sarà loro compito accettarla per renderla visibile tra i contenuti recenti visualizzati nella *home-page* di DataRep.

The screenshot displays the INGV (Istituto Nazionale di Geofisica e Vulcanologia) Open Access Repository interface. At the top, the INGV logo and name are visible, along with a search bar and a 'Communities' link. Below the header, a banner indicates 'OE DataRep - Beta Version - Only for Evaluation and Testing'.

The main content area is divided into two sections:

- Recent uploads:** This section features a dataset entry titled 'Temporal evolution of local seismicity in the island of Vulcano, Italy, in 1988 (TESV_1988)'. The entry includes a 'View' button, a 'Dataset' label, and an 'Open Access' badge. The author is listed as Falsaperla, Susanna. A brief description states: 'The dataset covers the time span from January 1 to December 31, 1988. It encompasses three classes of seismic events recorded on the island of Vulcano, in Italy. Here the local seismicity has varying origins and characteristics, from seismic events related to fluid dynamics to brittle failure...'. The upload date is 'January 26, 2022'. A 'More' button is located below the description.
- Tweets by @INGVvulcani:** This section shows a feed of tweets from the official INGV Twitter account. The top tweet, dated January 25, 2022, announces the publication of a weekly bulletin for the monitoring of volcanic, geochemical, and seismic activity at Vulcano. It includes a line graph titled 'daily bul 1922 Bul' showing data from 1950 to 2022. The graph plots '1922 Bul (Hz)' on the y-axis (ranging from 0 to 300) against time on the x-axis. The bottom tweet, dated January 25, 2022, announces the publication of a weekly bulletin for the monitoring of volcanic, geochemical, and seismic activity at Etna. It includes a link to the bulletin and the hashtag #sorveglianzavulcanica.

The footer of the page contains navigation links for 'About', 'Blog', 'Help', 'Developers', and 'Contribute'. The 'About' section includes links to 'About', 'Policies', 'Infrastructure', 'Principles', and 'Contact'. The 'Help' section includes links to 'FAQ', 'Features', and 'Support'. The 'Developers' section includes links to 'REST API' and 'D4P-PWH'. The 'Contribute' section includes links to 'GitHub' and 'Donate'. The footer also includes the text 'Powered by CERN Data Centre & Invenio.' and links for 'Status', 'Privacy policy', 'Terms of Use', and 'Support'.