EUROVOLC

European Network of Observatories and Research Infrastructure for Volcanology

Deliverable Report

D5.3: EPOS geochemical database participation: Integration of geochemical dataset of volcanic gases in atmosphere not implemented in EPOS-IP

Work Package:	*Consolidation of geochemical gas monitoring across VOs*	
Work Package number:	*WP5*	
Work Package leader:	*INGV*	
Task (Activity) name:	*Integration of geochemical dataset of volcanic gases in atmosphere not implemented in EPOS-IP.*	
Task number:	*Task 5.3*	
Activity leader	*INGV (Cinzia Federico)*	
Authors:	*F. Grassa, F. Viveiros, B. McKormick-Kilbride, M. Bonifaci, E. Ilynskaya, I. Iribarren, N. Luengo, L. Moreno, E. Manson, S. Moune, M.A. Pfeffer, C. Silva, M. Burton, S. Caliro, A. Di Muro, B. Esse, P. Labazuy, M. Liotta, M. Liuzzo, R. Moretti, A. Murphy, G. Salerno, P. Torres and A. Varna, L. Cacciola, G. Messina*	
Type of Deliverable:	*Report [x]* *Prototype [ ]*	*Demonstrator [ ]* *Other [ ]*
Dissemination level:	*Public [x]* *Prog. Participants [ ]*	*Restricted Designated Group [ ]* *Confidential (consortium) [ ]*

Programme: H2020
Project number: 731070

Contents

Summary

We report on the activities performed within Task 5.3, focused on the harmonization and integration of data and related information on volcanic gases managed by the European volcanological community. The final target is to propose a structure of database for geochemical data, obtained from field surveys with simultaneous measurements or with subsequent laboratory analyses.

The work of harmonization benefited from the joint field survey performed in February 2019 at the fumarolic field of the volcano Furnas (Azores), when the different participants compared their methodologies of gas sampling and discussed the parameters which better characterise data and methodologies. The work of Task 5.3 has been harmonized with the task finalized to the definition of best practices (EUROVOLC Task 5.1 and 5.2 reports).

Full disclosure of details regarding the sampling site, the sample characteristics, the sampling and analytical processes, is required for further data analysis, interpretation and comparison.

In particular, some procedures of sampling and analysis of gas and vapor can be quite complex, and many details are required for the full reproducibility of the data.

Indeed, the procedures of sampling and analysis of fumarolic vapor revealed some key features that can be potential sources of variability, such as sampling apparatus, duration of storage before analysis, the analytical procedure, the data treatment, among others.

The activities performed by the participants to the task are summarised below:

i) the first meeting was held in February 2019, during the second EUROVOLC Annual Meeting, in Sao Miguel (Azores). During this meeting, the participants agreed with the general structure of metadata, already developed in the EPOS project, to be further implemented for the geochemical data obtained during the field campaign performed at Furnas volcano; given the work in progress being performed by IT people from INGV-OE for the EPOS web services, the task's participants agreed to commit to INGV the implementation of the database for data obtained during the Furnas field survey.

ii) during a second meeting, held in Catania (Italy) in January 2020, we discussed the metadata to be associated with the data collected during the field campaign performed at Furnas volcano;

iii) during web meetings (held in May 2020 and March 2021), the structure of the database, under implementation by the INGV personnel, has been presented;

iv) the information gathered by the participants to Task 5.1, concerning methodologies, materials, analytical procedures, etc, were implemented as metadata in the database;

v) several restricted web meetings were devoted to the matching of scientific requirements and technical solutions, finalized to the development of the database and the web services.

1. Introduction

With the advances in analytical chemistry and the development of techniques of sampling and measurements in the different geochemical spheres (lithosphere, atmosphere, hydrosphere, pedosphere), the amount of data has drastically grown in a few decades. At the same time, some high profile journals, having page restrictions, do not host the dataset, which are published just in graphical form. This requires an organization of data based on a 3D structure, e.g. an additional dimension with respect to a 2D spreadsheet or table. Indeed, a wealth of additional information is required to make geochemical data reproducible and reusable, and broaden the use of applied methodologies of collection and analysis. This kind of architecture underlies the relational databases.
In the last decades, some databases were developed in specific fields of investigation.
Various database exist for igneous rocks from ocean islands, island arcs, continents (GEOROC), mid-ocean ridges, back-arc basins (PetDB), the western United States, British Columbia, northern Mexico (NAVDAT), which are part of EARTHCHEM portal (http://www.earthchem.org/).

A relational database has been created for the compilation of certified values of reference material (GEOREM), which are a significant part of the traceability and reproducibility of geochemical measurements (Jochum et al, 2005). The metadata includes the uncertainty, uncertainty type, method and laboratory for the analytical data, sample information, and references.

The GEOTRACES database gathers data and data products on the marine biogeochemical, trace elements and isotopes, obtained from oceanic cruises performed by a variety of nations.
The management of data of the GEOTRACES program is performed by the GEOTRACES International Data Assembly Centre (GDAC), which promotes data sharing and collaboration between research groups.

The need for standardizing metadata and models in geochemistry fostered the creation of a committee for the Geochemical Earth Reference Model (GERM). The database working group (established during the 2001 GERM Workshop) has put together a proposal for Metadata Standard in Geochemistry, published by Staudigel et al. 2003. However, the Geochemical Earth Reference Model has been implemented only for some geochemical reservoirs, namely the Solid Earth and Seawater (https://earthref.org/GERMRD/reservoirs/).

Within the EPOS Geochemistry Working Group, the partners developed an embryonic structure of a geochemical database, mostly dealing with data acquired at high frequency in soil emissions and in volcanic aquifers.
In this report, we describe the design principles employed in creating a comprehensive conceptual schema for a geochemistry database, fitting the wide range of geochemical data, obtained by in situ measurements and/or laboratory analyses. As a starting point, we focus on data obtained under the umbrella of the EUROVOLC Project. The proposed structure can be implemented to fit other geochemical data, on groundwater, marine water, soil, atmosphere, i.e., in any geological setting. Inspired from the work of Staudigel et al. (2003), we tried to develop a common framework for any kind of geochemical datum (rock and fluid geochemistry) obtained in the laboratories, by introducing specific metadata for the different sample types.
Full disclosure of details regarding the sampling site, the sample characteristics, the sampling and analytical processes, is required for further data analysis, interpretation and comparison.
In particular, some procedures of sampling and analysis of gas and vapor can be quite complex, and many details are required for the full reproducibility of the data. Chemical elements can be determined by a variety of techniques and the results can be different depending on the method and instruments, because they may have different detection limits, bias or interference. The traceability of geochemical measurements requires the disclosure of all the analytical steps and relative instrumental uncertainty and, hopefully, the uncertainty of the whole analytical process.

Usually, all information concerning sampling and analytical procedures are described in a specific section in scientific articles, which generally includes the dataset. Unfortunately, the descriptions are sometimes incomplete and possibly refer to a general methodology, reported in cited papers or in data table subscripts.
Given the increasing need to organize data series or campaigns in structured databases, we propose to include all information, described so far within the scientific articles, as mandatory metadata, associated with either samples or every analyte.
During the implementation of the geochemistry database, we defined a series of metadata required to fully describe the geochemical data. We followed a conceptual pathway moving from the sampling site, through the sample collection, to the analyses in the laboratories. Each phase of the process of data production is associated with different classes of metadata. Among these metadata, we selected some of them as keywords to search and reach a single datum or a dataset.
In the following, we illustrate the conceptual model for the relational database and describe every metadata class, by also providing some examples from the presented case studies. Some further technical details will be available in the Deliverable D20.1 provided by the WP20.

2. Activities within the WP 5 in EUROVOLC

The work performed within the EUROVOLC Project, aimed at the definition of metadata for geochemical measurements and analyses of volcanic fluids and described in this report, has been harmonized with the tasks finalized to the definition of best practices (EUROVOLC task 5.1 and 5.2 reports). The survey was carried out on the 21st and 22nd February 2019 at Furnas volcano (Sao Miguel, Azores). The sampling of fumarolic vapor has been performed by five teams (CIVISA, CSIC, IMO, INGV, IPGP), through direct sampling according to the Giggenbach's procedure, applied in worldwide volcano observatories for volcanic monitoring.
The procedures of sampling and analysis of fumarolic vapor revealed some key features that can be potential sources of variability, such as sampling apparatus, duration of storage before analysis, the analytical procedure, the data treatment, among others.
Additionally, some teams (IPGP-OVSG, INGV, IPGP-OVPF, CIVISA, IMO) also performed in-plume measurements with MultiGAS-type instruments. This kind of measurement is usually performed with specific sensors, depending on the gas species to be detected. The type of sensors, their sensitivity, their lifetime, and atmospheric conditions (i.e. temperature and pressure) appear as key factors that can potentially influence the final results.
We focused on the development of a structure of a geochemical database that could host, in a first approach, the data obtained during the EUROVOLC field survey held at Furnas volcano, leaving open the possibility to implement the service with other type of geochemical data, with a particular focus to the volcanological community.

3. The architecture of the project

The project design has been developed into four main steps: i) an initial requirements analysis, ii) a successive conceptual model implementation for the data representation, iii) the implementation and testing of the database structure and iv) the design of functionalities needed to the users to interact with the data structure.

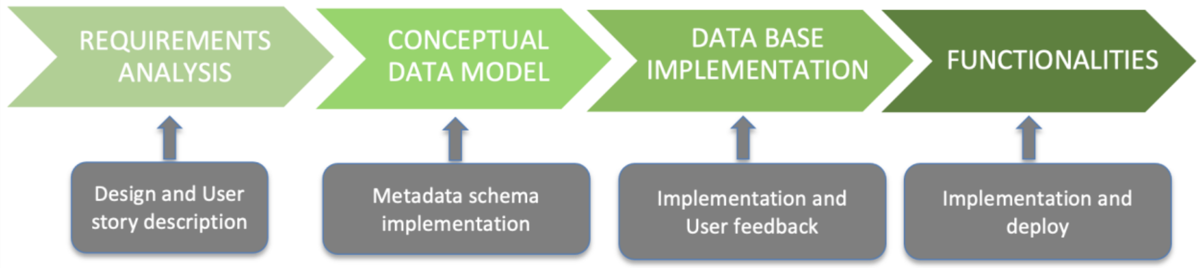


Figure 1- Steps of the project design

3.1 Requirements analysis

As a first step, we analyzed the needs of the volcanological community related to the collection, representation and organization of geochemical data on fluids. The result of this analysis was the creation of the *user stories* that generically describe the expectations of the users and how a generic user wants to interact with the system.

Table 1: User story description.

User Type	Functionality	User story process
Data provider	Sample information and geochemical analysis manage	• Collection of specific information on the samples (site, contact person, physical archive) • Organization of the parameters measured in the analysis of geochemical data • Supply of information on applied analysis techniques (accuracy, laboratory, tools)
Generic User	Data access, filtering, and discovery	• Access to efficient data classification, • Research on data through the application of filters • Data download

The table shows that two different types of users were defined:

1. *Data provider user*: generally speaking a single scientist, part of a scientific community, skilled in the acquisition/validation/publication of specific datasets (typical examples are a scientist who provide data; a technician, who deals with the analysis and organization of the obtained results; a responsible of an analysis laboratory and/or a project)
2. *Generic user*: able to discover and access the data, carry out searches by applying filters based on appropriately chosen keys, and finally download the data.

After a first round of iteration, specific tables have been produced by the geochemical/tech community (the working team) and a basic list of requirements have been released.

In particular:
- a generic user can use time and space to discover data;
- the system should be able to identify multiple analysis
- advanced filtering methods should be applied.

At the end of this phase, a first embryonic structure of the *geochem database* has been developed in order to be validated by the geochemical experts within the working team. This activity put in evidence many issues solved, step by step by IT specialists; new features have been proposed to enrich and complete the initial conceptual data model described, in its mature form, in the next section.

3.2 Conceptual data model

At the beginning of the task, a huge effort was expressed by scientific and technical staff in order to translate the geochemical community needs (within the EUROVOLC partners of the WP5) into an exhaustive scheme consisting of main macro levels chosen by the community. It is the basis concept that provides a representation of the *metadata structure* for geochemical data.

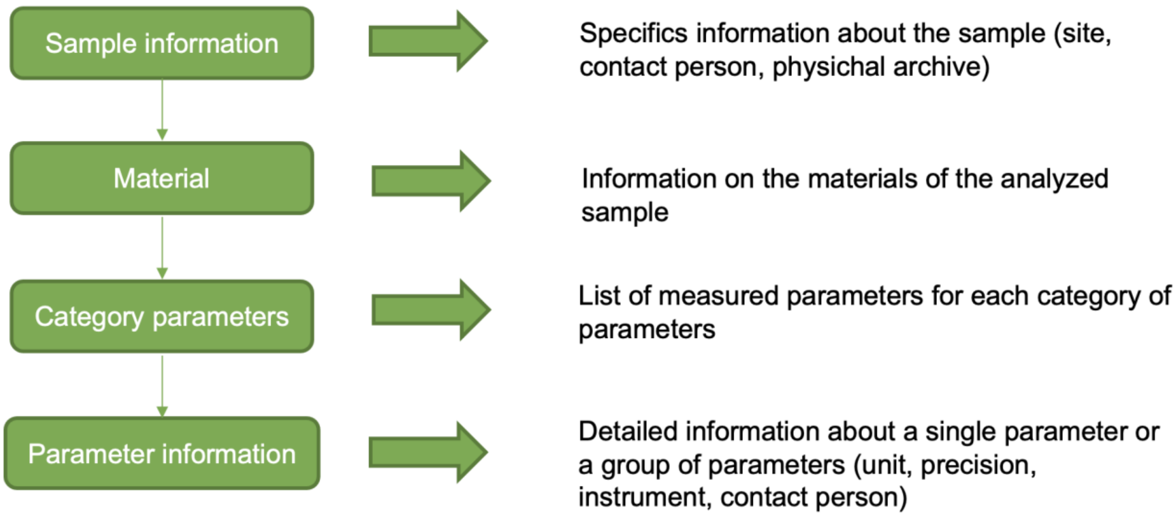


Figure 2 - Macro-levels for the metadata structure.

In detail, the first level contains the information related to *in situ* measurements, the sampling site and its procedure, information about the collector and the physical archive. The other levels include a list of metadata related to the sample treatment and related analysis.

In particular, the second level gathers the information about the analyzed material, which means the single phases, which will undergo specific treatments and analyses in the laboratories, after collection in specific containers, or measured with specific techniques in the field.

The third level represents the groups of parameters sharing some affinities, concerning, for instance, the analytical procedure or the instrument used for their determination.

The fourth level gathers the metadata related to every analyzed parameter, including measure units, analytical procedure and instrument, precision and accuracy.

3.3 Database implementation

Based on the four levels of the conceptual data model, described in section 3.2, the structure of a relational database has been proposed and a prototype of the main queries have been submitted to the geochemical experts, according to the general structure represented in Figure 2. The workflow of the specific metadata referred to the main different classes are listed below.

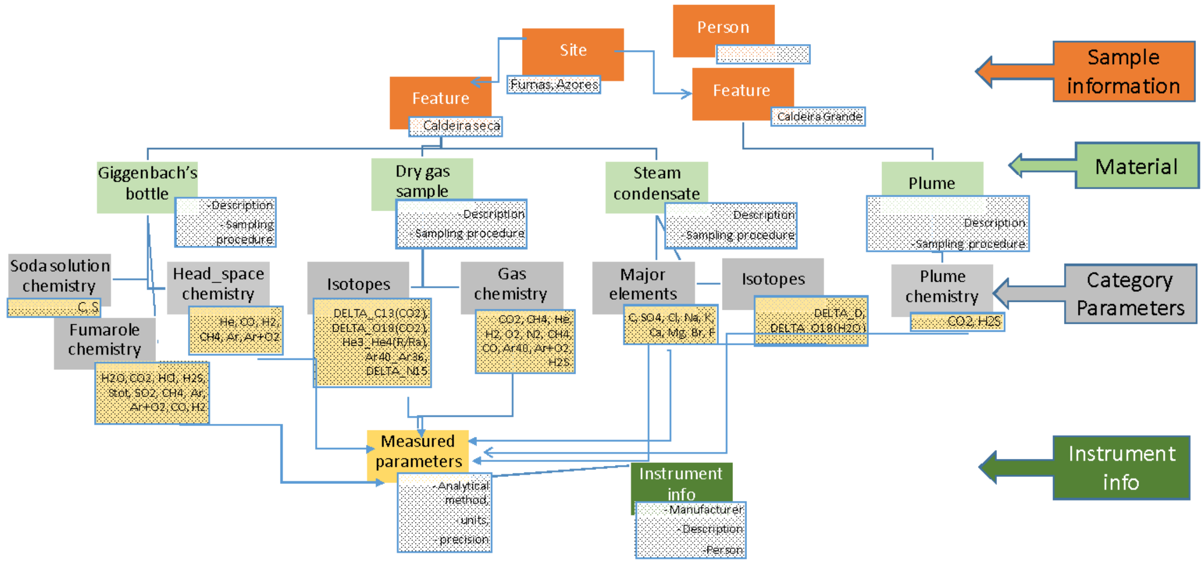


Figure 3 - Metadata structure

3.3.1 Site and feature classes

The metadata associated with the Site class are reported in Table 1.
The site of collection or measurement can be a single point, identified by latitude and longitude, or a volcano, identified by a Global Volcanism Program ID or the European Catalogue of Volcanoes ID, or an area, defined by the coordinates of a bounding box.
The "site" is related directly to the "material" class, or to the "feature" class.

Table 2: Metadata related to the “Site” class.

***”site”* class**		
URI	**Usage note**	**Example**
name	Name of the site	Furnas
GVP	GVP* Global Volcanism Program, Smithsonian Institution. https://volcano.si.edu/search_volcano.cfm	
country	Country where the site is located	Portugal, Azores
site_lat	Latitude of the site	
site_lon	Longitude of the site	
volcanic_area	Name of the volcanic area	Furnas
area_longW	If the site is an area (box), this is the Longitude W of the box	-25.3619
area_longE	If the site is an area (box), this is the Longitude E of the box	-25.2883
area_latS	If the site is an area (box), this is the Latitude Sud of the box	37.747°
area_latN	If the site is an area (box), this is the Latitude Nord of the box	37.786°

Indeed, the site of measurement and/or sample collection can be further detailed by the description of the natural or man-made feature under study. A large variety of features are possible: a fumarole, a mofete, a lava flow, a drilled well, a spring, the position of in-plume measurements, etc. When the feature is described, the site class represents a more generic attribute (the volcano, for instance). The class feature should contain as much information as possible about the sampling site, including pictures, sketches, and geometric details.

Table 3: Metadata related to the "Feature" class.

***"feature"* class**		
URI	**Usage note**	**Example**
feature_name	Name of the geochemical feature studied during a field survey or node of a permanent sampling network.	Caldeira Seca
sample_type	It can be a well, a spring, a fumarole, a soil gas sampling site, atmosphere, rock sample	Fumarole
sample_lat	latitude of the sampling point	37.772694°
sample_long	latitude of the sampling point	-25.304224°
sample_elevation	elevation of the sampling point	201 m
sample_collector	Name of the person who collected the samples	Fausto Grassa
Sample_photo_link	Link to the picture of the sampled feature	
survey_or_network_info-url	Information about the field campaign, the permanent network, project url or information page link	Eurovolc First survey Task 5.1

The site and the feature classes are associated with name, coordinates, elevation, and the references to a person (collector, responsible for the measurement campaigns, etc.).

3.3.2 Material class

The class "Material" represents the physical object that is collected in the field and that will undergo chemical analysis in the laboratories, or even the phase collected or measured (the fumarolic steam, the dry gas, the fumarolic plume). In a single site or feature, different samples and sample types are devoted to a variety of analyses. For example, in a fumarole, the vapor is collected for chemical analysis in typical two-way or single-way bottles, partially filled with a soda solution, according to the so-called Giggebach's method (Giggenbach, 1975). For the analysis of the vapor stable isotopes, the water vapor is condensed within glass or plastic vials or bottles, whereas the dry gas is usually collected in glass tubes, equipped with one or two stopcocks. As a result of the Task 5.1, concerning the best practices for fumarole sampling (EUROVOLC task 5.1 reports), the characteristics of the sampling devices (volume, type of stopcocks), the amount of basic solution where acidic gases are condensed have been demonstrated to affect the sample preservation and data quality. The detailed description of the materials used for the sample storage is required. Additionally, the date and (eventually) the duration of the collection of a given material provide clues on the good preservation of samples before the analyses in the laboratories, because the chemical composition of a fluid sample can be modified by the

different physical conditions encountered in the shallow environment, and these variations are time-dependent.
For the analysis of plume composition, different methods can applied, namely (i) remote sensing (e.g. scanning and imaging UV spectrometers, open path FTIR), (ii) in situ measurements (e.g. MultiGAS-type instruments) and (iii) gas or aerosol sampling for later laboratory analysis (e.g. filter packs). Each methodology requires a specific and standardized procedure, discussed by the participants in Task 5.2. These details are to be listed as metadata in the material class ("sampling_procedure" and "description").

Table 4: Metadata related to the "Material" class.

***"material"* class**		
URI	**Usage note**	**Example**
sampling_procedure	Procedure for sample collection and/or measurement	Giggenbach's method with funnel
description	Description of the collected and/or measured material	Giggenbach's bottle, NaOH 4M, volume bottle = 154.32 ml P head space = 303 mbar density=1.11 g/ml Bottle weight=145.87 g Weight bottle+soda solution = 203.665 g, Weight bottle+condensate=246.597 g, Weight solution extracted in lab=99.9 g 2 Torion stopcocks
comment		
material_label	Name of the material	PI3
sampling_date	Date of sample collection and/or measurement	21/02/2019
sampling_start	Start time of sampling	11:20
sampling_end	End time of sampling	11:31

3.3.3 "Category parameters" class

This class includes the groups of parameters that can be determined in each material. These categories have to be considered as reservoirs of parameters, which can contain any parameter having specific geochemical attributes. For instance, in the material "dry gas" the proposed categories of parameters that can be determined are "Gas chemistry", which includes the gaseous species analysed in gas samples (CO_2, O_2, N_2, CH_2), and "Isotopes", which groups the isotope composition of some elements. In the material "Giggenbach's bottle", the proposed categories of parameters are the "soda chemistry", which groups the parameters that are determined in the soda solution (usually C, S-bearing species, F^-, Cl^-), the "head-space chemistry", which includes the gas species non-condensed in the soda solution and enriched in the head-space of the Giggenbach's bottles. An additional "category parameter", referred to the material "Giggenbach's bottle", is the category "fumarole chemistry", which includes the parameters retrieved from the analyses of the soda solution and the head-space chemistry (plus some additional parameters, such as the volume and pressure of the head space, the volume and density of the soda solution), and represents the actual composition of the fumarolic vapor. In the material "steam condensate", the categories of parameters which we propose are "isotopes", which include water stable

isotopes (δD and $\delta^{18}O$), “major elements” and “trace elements”, which group any element of interest, generally occurring in concentrations of mg or ug per liter, respectively. Even in this case, these reservoirs can be filled with any parameter of interest.
The “Category parameters” class is related to
- the list of parameters to which it refers;
- the “material” class.

3.3.4 “Parameters” class

This class contains any parameter (element, compound, isotopic ratio, physico-chemical parameter) that can be measured in the field or in the laboratories. The single parameter can be included or not in a category of parameters. The metadata related to every parameter are

- the name of the parameter, referred to a vocabulary, possibly acknowledged by the scientific community; in this work, we adopted the vocabulary proposed by Earthchem (http://www.earthchem.org/);
- the measure units;
- the analytical precision;
- the analytical procedure, including eventual treatments before the analyses and the measurement technique;
- the instrument used for the analyses;
- the detection limit;
- the calibration type and the range of concentration of the standards;

Table 5: Metadata related to the “Parameter” class.

”parameter” class		
URI	**Usage note**	**Example**
parameter_name	Name of the measured parameter (chemical element or compound, isotopic ratio)	He
unit	measure unit	ppmv
precision	laboratory analytical precision	5%
date	Date of analysis	
laboratory	Reference of the Laboratory where the parameter is determined	INGV-Palermo
analytical_procedure	Analytical method and procedure	Gas chromatography. Glass bulb connected to the GC intro system and re-equilibrated to atmospheric pressure before analysis
detection limit	The lower concentration that can be measured	0.0005%
standard_range	The range of concentration of the used standards	0.0424% 0.0101% 0.00103%
calibration_type	type of calibration (linear, exponential…)	linear

dataset	Dataset containing the parameter	Eurovolc First survey Task 5.1 fumarole analysis
references	If the datum is already published, it is the doi	

Each parameter is related to
-the "Category parameters" class
- the "Instrument_info" class.
Some parameters (i.e. fumarolic temperature) are measured in the field and, therefore, they are directly related to the "feature" class.
Some parameters could be part of a dataset, related to a field campaign, or a given temporal interval.

3.3.5 "Instrument_info" class

This class contains all information about the instruments used in the field and in the laboratories. The instrument is identified by the manufacturer and model, and should be associated with the laboratory where it is installed or stored, and the reference person, responsible for its use and maintenance.

Table 6: Metadata related to the "Instrument_info" class.

"instrument_info" **class**		
URI	**Usage note**	**Example**
instrument_name	commercial name of the instrument	7890B
manufacture	Manufacturer of the instrument	Agilent
description	Description of the instrument	Gas-chromatograph equipped with Poraplot U 25m x 0.53 mm and Molsieve 5A 25m x 0.53 mm columns. fluxed by Ar, detectors TDC and FID with methanizer
laboratory	Reference of the Laboratory where the instrument is located	INGV-Palermo

3.3.6 "Person" class

The class "person" lists any relevant person involved in the project, in the field or laboratory activities. Consequently, this class can be associated with various classes. It lists the information related to the responsible of the sampling campaign, to the sample collector, to the responsibles of the laboratories, to the analysts. Here below, two examples of person classes used in EUROVOLC.

Table 7: Metadata related to the "Person" class.

"person" **class**		
URI	**Usage note**	**Example**
person_type	Role of the person	Responsible of work package
family_name	Person's family name	Grassa
given_name	Person's given name	Fausto
nationality	Person's nationality	Italian
institution	Person's place of work	INGV-PA
email	Person's email address	fausto.grassa@ingv.it

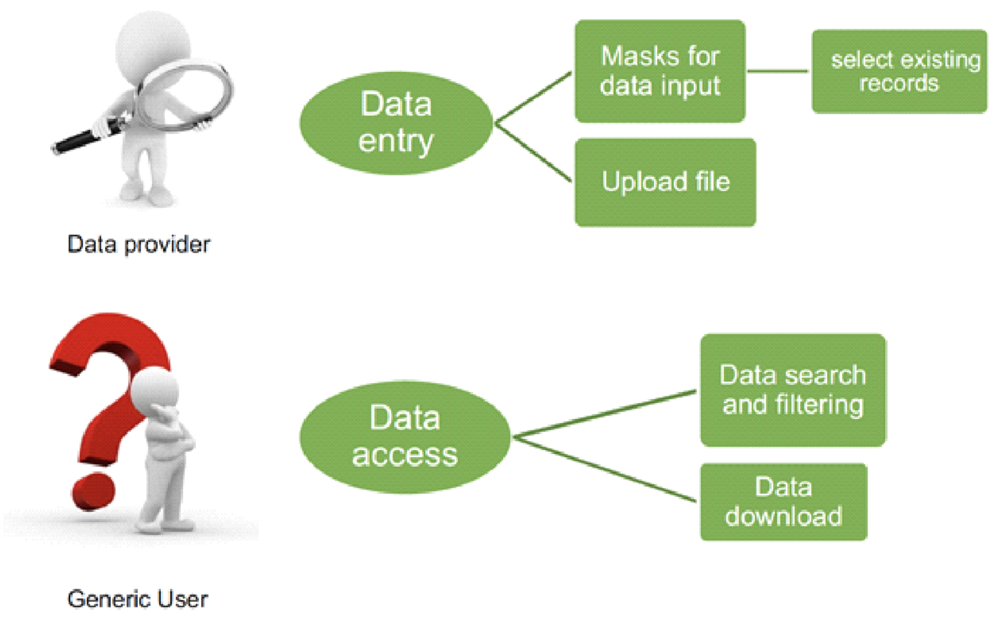
"person" class		
URI	**Usage note**	**Example**
person_type	Role of the person	Responsible for the laboratory of Gas chromatography
family_name	Person's family name	Longo
given_name	Person's given name	Manfredi
nationality	Person's nationality	Italian
institution	Person's place of work	INGV-PA
email	Person's email address	manfredi.longo@ingv.it

3.4 Functionalities

Taking into account the work done in the R*equirements analysis* section (see section 3.1), the tech team proposed an internal validation phase, in order to verify the development status, compared with the proposed deliverables. Some tests have been proposed: chosen datasets have been ingested by the system and located into the database structure, tables were verified/improved, and the capability of the database to execute proper queries was tested. At the end of the process, a mockup design of a custom searching data portal was proposed to choose the better way to show the results.

After creating the database structure and testing its correctness, we implemented the functionalities to interact with it. According to the different types of users considered, it is possible to distinguish:

- the data provider functionalities, to insert and manage the data ;
- the generic user functionalities, to discover and access the data.



The data providers can use two different tools to enter the data:

- entry masks, developed in PHP, with the possibility of selecting records present in the database;

- the possibility of inserting information into the database by uploading structured files (excel, csv, json).

The generic user functionalities (generic user who can be a researcher, a student or any type of user who wants to interact with the system) have been developed:

- with the creation of search masks, implementing REST web services in Python, which allow to filter the data by some specific keywords;
- by allowing the data download.

4. To populate the database

To gather the information required to populate the metadata structure of the database, we referred to the queries, formulated by the task leader of the Task 5.1 and shared among the participants, and the annotations recorded during the field campaign at Furnas volcano.

During the phase of sampling, the partners used glass bottles of different capacity and different stopcocks, and different amounts of NaOH solution, although they used the same sampling line. Beyond the natural temporal variability of the natural system, some variables were expected to affect the final results and, with this in mind, some additional tests were performed. The tests, performed in July and September 2020 by CIVISA and in August 2020 by IPGP-OVSG teams, were aimed at:

- evaluate the impact of different amounts of soda in the bottles (controlling the volume of the headspace);
- check the effect of time between sampling and analyses of the gases collected.

According to the inferences obtained after these tests and the long-term expertise of the involved partners in the collection of fumarolic gases, we proposed to include in the description of the collected "materials" and, specifically, for the Giggenbach's bottle, the information concerning

- the volume of the bottle
- the amount of soda solution
- the amount of collected vapor
- the date of sampling
- the duration of the sample collection
- the date of analysis.

Some of these data, among others, are also required to compute the composition of the fumarolic vapour, after the analysis of the soda solution and the gas species in the bottle headspace.

The information related to every analyzed parameter was gathered through the query, represented in Table 8.

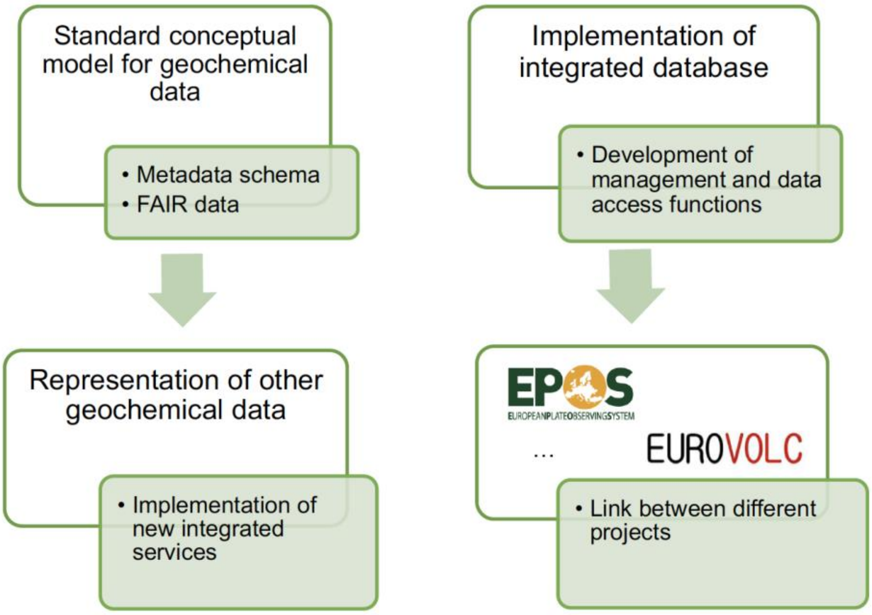
The table 1 lists the queries concerning the analytical facilities of some of the partners (CIVISA, INGV-PA, INGV-OV, IPGP-OVSG), including the analytical procedures, the technique and the instruments used for each gas species or isotope, the accuracy and limits of detection, the calibration range.

Table 8: Query about the analytical facilities.

Questions - TASK 5.1 WP5																
1. Please identify your laboratory																
1.1. Name of the responsible for the analytical procedures																
1.2. Name of the responsible of the laboratory (and that confirms the analyses)																
ANALYTICAL PHASE	H_2O	CO_2	H_2S	^{40}Ar	O_2	N_2	CH_4	H_2	He	CO	$^{40}Ar/^{36}Ar$	$\delta^{15}N$	$\delta^{13}C_{CO2}$	$\delta^{18}O_{CO2}$	$\delta^{18}O_{H2O}$	$\delta^{2}H_{H2O}$
2. Based on the above gas species, which method (Chromatography, titration...) you use to analyse each of the gas species?																
3. Name the instrument used to analyse each gas specie																
3.1. Refer the brand of the instrument referred on point 3.																
3.2. Refer the model of the instrument																
3.3. Refer the accuracy of the measurement (for each specie), if known																
3.4. Refer the limit of detection/quantification of the measurement (for each specie), if known																
4. Which standard you use for each of the gas specie (for instance, gas bottles with a specific mixture of gases that should be mentioned; known concentration of gas in a solution...)																
5. Is this the first time you perform inter-laboratorial comparison of analyses?																
6. Please share bibliographic references where the methodology used in your laboratory is described.																

Concerning in-plume measurements, participants in Task 5.2 gathered the information on sensors, raw gas concentrations and intervals of data where the gas ratios were computed.

5. System operational possibilities and future perspectives



Once a mature version of the geochemical database has been reached and the internal validation passed, the best way to check the system is to choose a team of external experts that might check functionalities suggesting improvements and reporting bugs. A survey among them could be the best way to collect precious feedback, very useful to improve the whole system.

The goal of the proposed work is to give the geochemical community a tool, rich in features and dynamic searches for general purpose uses. Concerning future perspectives, different types of analysis and samples can be added into the structure, simply re-using the proposed metadata structure.

For these reasons, the metadata structure developed in Task 5.3 has a double role: i) to describe accurately the entire experimental process and ii) to make the data findable and accessible to the scientific community.
In this way results can be exposed in terms of FAIR data accessible by a proper Web Data Portal and/or by using Rest Services, largely used by machine-readable processes. Moreover, many projects worldwide require that scientific communities are mature enough to improve their own data standardization by using the so-called "intelligent data format". The Volcanological community of the EUROVOLC project is, for sure, the best field for developing these aspects having an important impact in terms of sustainability because of its own strict connection with big Research Infrastructures like EPOS ERIC.

6. Main Considerations

The measurements and sample collection in fumaroles and in fumarolic plumes are time consuming and foresee various steps from sampling to analysis and data elaboration, which require standardized procedures. We recognize that the different steps of the whole process should be described, and opportune metadata have to be implemented in data services to make data FAIR (findable, accessible, interoperable, and reusable). Additionally, the experimental reproducibility of geochemical data relies on the accurate description of all the steps leading to their determination.
We used the field campaign performed in Sao Miguel (Azores, Portugal) in February 2019 as a case study, and defined the metadata structure tailored on fumarole data. Although the gathering of all the information from the different partners is not fully achieved, the service will be shortly operative with a user interface for data and metadata entry.
Additionally, the metadata structure allows the implementation of the service with other categories of geochemical data, obtained by either sampling and analysis in the laboratories or by in situ measurements. It is specifically developed to host the data provided by the volcanological observatories, but its structure can be implemented for geochemical data obtained in different geological settings.

7. References

Giggenbach, W.F. (1975) – A simple method for the collection and analysis of volcanic gas samples. Bulletin of Volcanology, 39, 132–145.

Jochum, K. P., Nohl, U., Herwig, K., Lammel, E., Stoll, B., & Hofmann, A. W. (2005). GeoReM: a new geochemical database for reference materials and isotopic standards. Geostandards and Geoanalytical Research, 29(3), 333–338.

Staudigel, H., Helly, J., Koppers, A. A., Shaw, H. F., McDonough, W. F., Hofmann, A. W., ... & Zindler, A. (2003). Electronic data publication in geochemistry. Geochemistry, Geophysics, Geosystems, 4(3).

L. Cacciola, G. Messina, D. Reitano, R.A. Corsaro, C. Federico, (2021) "GeoChem database: design and implementation of a data structure for geochemical data", https://osf.io/6dz2t/.

Participants

Magali Bonifacie - *IPGP* - *Institut de Physique du Globe de Paris*

Christophe Brunet - *IPGP* - *Institut de Physique du Globe de Paris*

Lucia Cacciola - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Stefano Caliro - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Andrea di Muro - *IPGP* - *Institut de Physique du Globe de Paris*

Franck Donnadieu - *UCA* - *Université Clermont Auvergne*

Edda Elisa Falcone - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Cinzia Federico - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Fausto Grassa - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Evgenia Ilyinskaya - *UNIVLEEDS* - *University of Leeds*

Ilazkiñe Iribarren - *CSIC* - *Agencia Estatal Consejo Superior de Investigaciones Científicas*

Philippe Labazuy - *UCA* - *Université Clermont Auvergne*

Frederic Lauret - *IPGP* - *Institut de Physique du Globe de Paris*

Marcello Liotta - *INGV* - *Instituto Nazioanale di Geofisica e Vulcanologia*

Marco Liuzzo - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Natividad Luongo - *CSIC* - *Agencia Estatal Consejo Superior de Investigaciones Científicas*

Daniela Matias - *CIVISA* - *Centro de Informação e Vigilância Sismovulcânica dos Açores*

Giuseppe Messina - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Lucia Moreno - *CIVISA* - *Centro de Informação e Vigilância Sismovulcânica dos Açores*

Roberto Moretti - *IPGP* - *Institut de Physique du Globe de Paris*

Severine Moune - *IPGP* - *Institut de Physique du Globe de Paris*

Sérgio Oliveira - *CIVISA* - *Centro de Informação e Vigilância Sismovulcânica dos Açores*

Melissa Pfeffer - *IMO* - *Icelandic Met Office*

Vincenzo Prano - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Danilo Reitano - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Vincent Robert - *IPGP* - *Institut de Physique du Globe de Paris*

Francesco Salerno - *INGV* - *Istituto Nazionale di Geofisica e Vulcanologia*

Catarina Silva - *CIVISA* - *Centro de Informação e Vigilância Sismovulcânica dos Açores*

Pedro Torres - *CSIC* - *Agencia Estatal Consejo Superior de Investigaciones Científicas*

Fátima Viveiros - *CIVISA* - *Centro de Informação e Vigilância Sismovulcânica dos Açores*