RÉSIF-SI: A Distributed Information System for French Seismological Data

Catherine Péquegnat *et al.*

Abstract

The Résif project, which started in 2008, aims at gathering under a common research infrastructure the French seismological, Global Navigation Satellite Systems and gravimeter permanent networks, as well as the mobile instrument pools. A central part of Résif is its seismological information system, Système d'Information de Résif (Résif-SI) (started in 2012), which is in charge of collecting, validating, archiving, and distributing seismological data and metadata from seven national centers. Résif-SI follows a distributed architecture, in which the six data collection and validation centers (A-nodes) send validated data and metadata to a national data center (Résif Data Center [Résif-DC]), which is the central point for data archiving and distribution. Résif-SI is based on international standard formats and protocols, and is fully integrated into European and international data exchange systems (European Integrated Data Archive, European Plate Observing System [EPOS], Incorporated Research Institutions for Seismology, International Federation of Digital Seismograph Networks). In this article, we present the organization of Résif-SI, the technical details of its implementation, and the catalog of services provided to the end users. The article is aimed both at seismologists, who want to discover and use Résif data, and at data center operators, who might be interested in the technical choices made in the implementation of Résif-SI. We believe that Résif-SI can be a model for other countries facing the problem of integrating different organizations into a centralized seismological information system.

Cite this article as Péquegnat, C., J. Schaeffer, C. Satriano, H. Pedersen, J. Touvier, J.-M. Saurel, M. Calvet, L. Stehly, P. Arnoul, P. Bollard, *et al.* (2021). RÉSIF-SI: A Distributed Information System for French Seismological Data, *Seismol. Res. Lett.* **92**, 1832–1853, doi: 10.1785/0220200392.

Introduction

Réseau sismologique et géodésique français – French seismological and geodetical network (Résif, see Data and Resources) started as a project in 2008, with the aim of grouping into a single research infrastructure all the major national and regional seismic, the Global Navigation Satellite Systems (GNSS), and gravimeter permanent networks and mobile pools in France. Hand in hand with this objective was to improve and rationalize the data distribution system. GNSS data from all Résif partners are distributed from the national Résif data center for GNSS data (Résif, 2017), which also hosts the EPOS GNSS data gateway, whereas, the Résif data from permanent gravimeters are directly integrated into dedicated international data centers (Voigt *et al.*, 2016). In this article, we focus on the seismological part of the Résif information system (Résif-SI), for which the architecture, technical implementation, and governance structure has proven to be adequate, to address data distribution through a coordinated national cooperative. The present article refers to seismic waveform data and metadata: Higher-level earthquake products in France are now also being integrated into Résif, as discussed in Masson *et al.* (2021).

Prior to Résif, French seismic waveform data were distributed from five institutions:

- Commissariat à l'énergie atomique et aux énergies alternatives (CEA), commissioned by the French ministry of the interior to establish earthquake alerts toward the national authorities, operated (and still operates) a dedicated seismic network with high-availability data transmission (Massinon and Plantet, 1976; Résif, 2018). Data from this network were available through direct cooperation with CEA.
- ISTerre (then Laboratoire de Géophysique Interne et Tectonophysique [LGIT]) in Grenoble distributed data from the national strong-motion network (RAP, Résif, 1995) via the RAP National Data Centre (RAP-NDC, Pequegnat *et al.*, 2008). ISTerre also ensured national distribution of data from temporary seismic deployments using the national instrument pool, but most research laboratories also ran their own, smaller instrument pools.
- Institut de physique du globe de Paris (IPGP) distributed data from the GEOSCOPE network (Institut de physique du globe de Paris [IPGP] and École et Observatoire des

Full author list and affiliations appear at the end of this article.
*Corresponding author: satriano@ipgp.fr

Sciences de la Terre [EOST], 1982), via the GEOSCOPE data center, with a historical strong role of GEOSCOPE in the promotion of open data distribution at a worldwide level (Roult *et al.*, 1999). IPGP also operates three seismic networks in the Antilles (IPGP, 2008a,d,b) and one in Reunion Island (IPGP, 2008c), and those datasets were at the time planned for integration in the IPGP data distribution.
- Géoazur and École et Observatoire des Sciences de la Terre (EOST): The metropolitan French broadband and short-period data were loosely integrated in a national cooperation related to long-term Earth observations, but each regional network was operated independently, with no joint instrumentation policy, and distributed mainly by these two institutions.

The principle of transparent access to data through the common international distribution tool NetDC (Casey, 1999) was agreed as a common goal, through an additional transparent layer to a virtual data center, Fédération de l'Observation Sismologique Française (FOSFORE) (Shapiro *et al.*, 2008). Practically, the national French broadband network was underperforming, for a lack of sufficient station coverage and homogeneous instrumentation policy. In addition, the difference in technical implementation of data and metadata distribution, data holdings and data center availability between the five locations meant that the federated approach, in practice, needed to evolve to a new system.

When Résif was created as a concept, it moved under the direct supervision of the French Ministry of higher education and research, and integrated 18 organizations into a national consortium, which was signed in 2011. As a first success of this coordination effort, Résif obtained the funding (9.3 M€, with, approximately, the same amount committed by the partners in terms of staff time) for a project with four main components: The construction of a national broadband network; the extension and renewal of the GNSS network, and of the different mobile pools; the creation of a national Résif information system. The constraints from labor policies and civil servants meant that the permanent staff would transfer without relocation from the old system to the new, with additional temporary staff provided to accompany the changes.

The result of this effort is presented in this article. The architecture of the Résif information system (French: système d'information, hereinafter, Résif-SI, re3data.org, 2016) is fully distributed, matching the distributed architecture of the French seismic networks, which relies on CEA and on regional operators (mainly observatories hosted at universities, and with strong Centre National de la Recherche Scientifique [CNRS] involvement) for production and validation of data and metadata. These operators, through the so called "A-nodes," have full responsibility of pushing data and metadata into the data distribution center, via an automated procedure that allows for regular updates and automatic replacement in the data bases, as instruments are replaced, errors (timing, instrument responses, etc.) are corrected, and data gaps are filled. The system is, at the same time, fully integrated, with a national management team (Executive Committee), a technical board, which meets monthly and which adopts joint technical solutions, a single national data distribution facility (hosted by the University of Grenoble), and a coordinated participation in international and European and collaborative instances, such as International Federation of Digital Seismograph Networks (FDSN, see Data and Resources), and Observatories and Research Facilities for European Seismology (ORFEUS; see Data and Resources). In particular, the national data distribution facility is a part of the European Integrated Data Archive (EIDA) federation, a service within ORFEUS that gives federated access to data from European data centers. The EIDA is described in Strollo *et al.* (2021).

Résif-SI has provided, within the constraints given, remarkable results, with a stable and high-quality data distribution and a full integration into the international collaborations. The data holdings and download statistics are on par with other large data centers in Europe.

We believe that the Résif-SI implementation can be a model for countries that face the challenge of regrouping observation networks and data distribution, and, particularly, not only those countries where the need for regional and institutional visibility and independence remain strong, but also where each institution alone does not have critical mass for running a high-availability data center.

The article is structured as follows: we start describing the overall Organizational Structure of Résif-SI, for then discussing the Data Management: Résif-DC system and, in particular, the technical choices behind the Résif-SI data center and how they have been functional to integrate Résif-SI into the international data exchange systems. The New Developments section illustrates four areas (large-N experiments, marine data, building and infrastructural monitoring, and a StationXML metadata editor) for which Résif-SI is at the forefront of current development in seismological data management. We conclude by presenting the upcoming challenges and the strategy to meet them in the Conclusion and Future Challenges section. The Data Access Details and Examples section is provided with details and practical examples on the different ways to access Résif-SI data.

Organizational Structure of Résif-SI

Résif-SI archives and publishes seismological data from 11 permanent networks and, approximately, 70 temporary networks (Fig. 1).

Seismological data and metadata are collected and validated by six centers run by Résif partners, which are called "A-nodes" (Fig. 2), which are also responsible to secure a copy of the data for two years. The data and metadata are then transmitted to the Résif seismological data center (Résif-DC, also called the "B-node") hosted by the University of Grenoble Alpes for archiving (storage and remote archive) and distribution (services, portal). Résif-DC is designed and operated by a technical

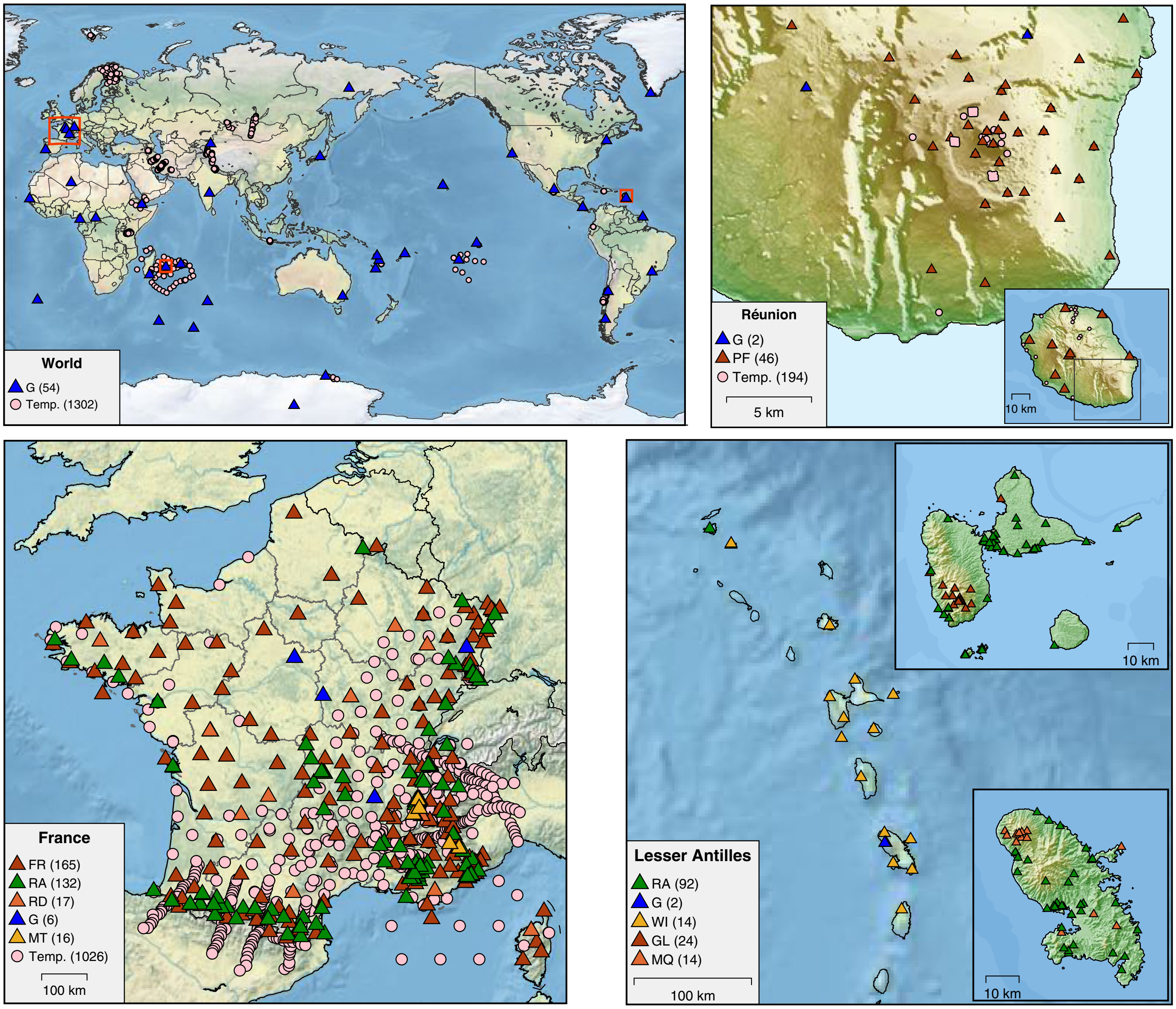


Figure 1. Seismic stations distributed by Réseau sismologique et géodésique français (Résif-SI) worldwide, in metropolitan France and in overseas Réunion and Lesser Antilles regions (location of these regions is indicated by the red boxes on the world map). Symbols and colors according to the network; number of stations for each network indicated in parentheses (see Table 1 for details on the network codes). The "Temp." label includes all temporary deployments (land mobile instrument pool SisMob and ocean-bottom seismometer instrument pool). The three pink boxes on the "Réunion" map are deployments of 100 sensors each (see Brenguier *et al.*, 2015). Note that the maps do not show all the Résif-SI data holdings worldwide, and other stations are available from permanent or temporary deployments in Europe, Chile, and Pacific Ocean. Note also that the maps include stations that are temporary or definitively closed. The color version of this figure is available only in the electronic edition.

team from ISTerre and from the Observatoire des Sciences de l'Univers de Grenoble (OSUG). Table 1 illustrates the typology of data available from each of the six operational A-nodes, whereas, Figure 3 shows the proportion of data archived at Résif-DC by type: accelerometric, velocimetric, and other. A seventh A-node, MARINE, operated by IPGP and Observatoire de la Côte d'Azur, is in its implementation phase, and will collect and validate the data from ocean-bottom seismometer (OBS).

Résif-DC, the national Résif data distribution center, is one of 19 global centers distributing data and metadata using formats and protocols that comply with FDSN (see Data and Resources) standards. It is also one of the 12 nodes of the EIDA (see Data and Resources).

From the very beginning of Résif-SI in 2012, it was defined that the quality of data and metadata, and their rapid and regular availability was the responsibility of the A-nodes, whereas the quality and continuity of data services depended on the national data center. This principle is still at the core of the

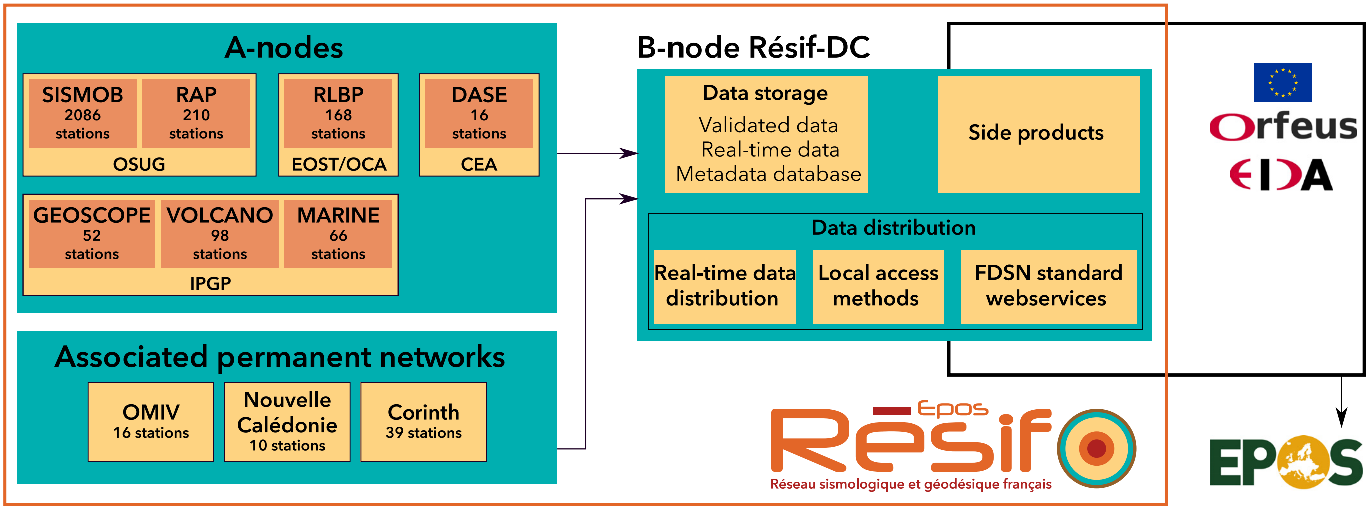


Figure 2. Structural architecture of Résif-SI. The A-nodes collect and validate data and metadata from permanent and temporary networks. They are also responsible to secure a copy of the data for two years. Data and metadata are then submitted to the B-node Résif-DC, which is responsible for long-term storage and distribution. Résif-DC also computes side products (e.g., power spectral density plots, McNamara, 2004), and is member of the European Integrated Data Archive (EIDA). Overall interoperability between data from different observations within a large part of Solid Earth in Europe is ensured by EPOS. The color version of this figure is available only in the electronic edition.

organization of Résif-SI, but all of Résif-SI jointly takes action to continuously improve data and metadata quality.

Data and metadata workflows

A-nodes take care of data validation and production of associated metadata for the observatories and instruments they are responsible of, according to their own workflows and objectives of completeness and quality.

Data integration is initiated, managed, and controlled by the A-nodes, which can add, modify, or replace data and metadata (the removal being under the control of Résif-DC) at their own discretion and, according to their rhythm, in an autonomous way. To this purpose, Résif-DC maintains the dedicated tool Résif data transfer (Volcke *et al.*, 2013). An important and explicit rule of Résif-SI is that Résif-DC does not in any way modify the data or metadata provided by A-nodes. The data and metadata completion and validation process at A-nodes is iterative. Because A-nodes are responsible for retaining all raw data collected by any means (including data in proprietary formats) for, at least, two years, all or part of the processing phases can be easily reversed, if necessary.

The integration protocol specifies:

- the formats allowed as input: miniSEED (Ahern *et al.*, 2012) or PH5 (Hess *et al.*, 2018) for data, StationXML 1.1 (International Federation of Digital Seismograph Networks [FDSN], 2019) for metadata;
- the transfer modalities—based on rsync (Tridgell and Mackerras, 1996) push and acknowledgment—based on rsync get or a webservice;
- the checks carried out before the metadata and the data are integrated into the database and the data archiving;
- the integration modalities: a product submission is a transaction;
- the structure of the integration report to the partner.

The integration controls are implemented at the data center level and have been jointly specified.

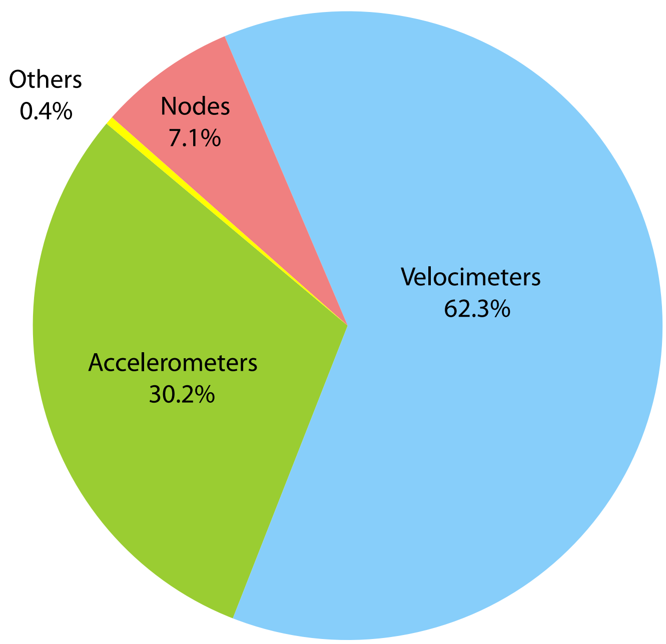


Figure 3. Distribution of data volumes in Résif archive, as of 1 January 2021. Velocimetric data represent 62.3%. Accelerometric data represent 30.2%, Data from very short-period "node" geophones represent 7.1% and other data, such as meteorological time series, accounts for 0.4%. The color version of this figure is available only in the electronic edition.

TABLE 1
Typology of Stations by Networks within Résif A-Nodes

Résif A-Node	FDSN Network Codes	Stations	NRT	BB	SP	SM	Others
RLBP	FR	167	145	167	6		3
	CL	39	15	37	19	4	3
	ND	10	1	10		7	
	MT	16	10	16	5		1
RAP	FR	49	39			49	
	RA	232	93			232	
GEOSCOPE	G	54	30	54		10	36
Volcano	WI	15	11	15		15	
	PF	46	36	29	20		
	GL	22	9	11	18		
	MQ	14	12	8	8		
SisMob	70 Temporary network codes	2482		1271	277	63	1022
CEA	RD	19	13	19		1	
	Total	3165	414	1637	353	381	1065

RLBP: French broadband permanent network (FR) cohosted by École et Observatoire des Sciences de la Terre (EOST) and Observatoire de la Côte d'Azur (OCA); also receives data from partner permanent networks (CL, ND, and MT). RAP: permanent accelerometric network (RA, FR), hosted by Observatoire des Sciences de l'Univers de Grenoble (OSUG). GEOSCOPE: GEOSCOPE global observatory (G), hosted by Institut de physique du globe de Paris (IPGP). Volcano: seismological instruments from the Observatories of Piton de la Fournaise (Reunion Island, PF), Guadeloupe and Martinique (GL, MQ, and WI), hosted by Institut de physique du globe de Paris (IPGP). SisMob: land mobile instruments, hosted by ISTerre; it also receives data from partner laboratories. Data can be embargoed up to 3 yr after the experiment start. Commissariat à l'énergie atomique et aux énergies alternatives (CEA): CEA broadband stations (RD). NRT: number of stations with, at least, one near-real time stream. BB: number of stations with, at least, one broadband channel (NRT or not). SP: number of stations with, at least, one short-period channel (NRT or not). SM: number of stations with, at least, one accelerometric channel (NRT or not). Others (NRT or not): tiltmeter, weather, hydrophone, wind, rotational sensor, mass position, etc. Note that for SisMob the "others" category includes very short-period "node" geophones (channel codes starting with "DP").

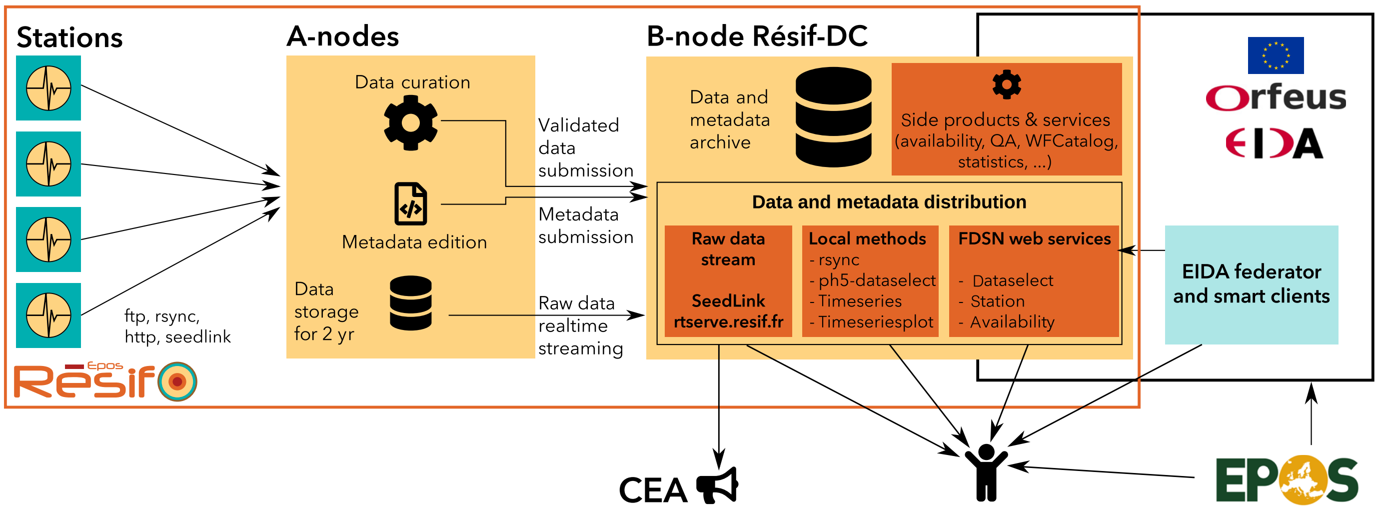
From the point of view of A-nodes, any product submission (simple or complex) is a transaction, whose characteristics is preserved and can be retrieved by a webservice. The integration of data and metadata is asynchronous, and the consistency between them is re-evaluated at the end of each transaction. The submission of not-yet-described or incompletely described data is allowed, to secure them as soon as possible and according to the needs of the A-nodes, but Résif-DC data services do not deliver any data whose metadata are not available. This workflow is illustrated in Figure 4.

From the point of view of Résif-DC, an integration of data or metadata is a succession of complex operations carried out by several independent workers (Fig. 5) that communicate with each other via an Advanced Message Queuing Protocol system (ISO/IEC 19464, 2014). Four databases are populated at this stage: (1) inventories of networks, stations, channels, and responses; (2) waveform inventories (PostgreSQL databases, PostgreSQL Global Development Group, 1996); (3) metrics and data quality information (MongoDB – EIDA WFCatalog database, MongoDB, Inc., 2009; Trani *et al.*, 2017); (4) elements for monitoring the integration operations (PostgreSQL). Consistency of information within and between databases (1), (2), and (3) is guaranteed by construction. Data are archived in a SeisComP Data Structure (SDS) structure (GFZ and gempa GmbH, 2008, if miniSEED data) or as HDF5 record (Folk *et al.*, 2011, if PH5 data).

Résif-DC also computes on the fly, with each integration, some useful side products. For instance, power spectral density (PSD) plots (McNamara, 2004) generated with PASSCAL Quick Look eXtended (PQLX) (McNamara and Boaz, 2010), but also availability plots generated with the help of a Standard for Exchange of Earthquake Data (SEED) data indexer (seedtree5, Volcke *et al.*, 2012).

Résif-DC exposes two webservices for A-nodes to manage their data and transactions:

- **http://ws.resif.fr/resifsi/transaction/1/**: to retrieve transaction records, for example, to analyze and correct rejected data or metadata (see Data and Resources);
- **http://ws.resif.fr/resifsi/orphanfile/1/**: to identify orphan data (without metadata) in the Résif archive for corrective action (see Data and Resources).



Other operations have to be requested through a help desk, because either they are rare, or there is no safe way to automate them (e.g., data removal).

Metadata enhancement

Résif-SI has progressively improved the quality and consistency of metadata by establishing a controlled vocabulary and recommendations for standardizing the content of text fields in metadata. Initially, the metadata of A-nodes were managed in local databases and formatted using the open-source portable data collection centers (PDCC) tool by the Incorporated Research Institutions for Seismology (IRIS, Casey, 2016) or in-house tools. Each A-node has its specificity regarding metadata. In particular, SisMob and RAP have to deal with a large diversity of instruments and very complex metadata. With this context, in-house tools to manage the metadata were mandatory. Indeed, all the constraints on the metadata for Résif-SI networks could not be included into any standard software (e.g., SeisComp, GeoForschungsZentrums [GFZ] and gempa GmbH, 2008) at the time.

Before 2019, metadata were edited and submitted in SEED dataless format (Ahern *et al.*, 2012), and Résif-DC ingested the metadata and delivered it in the normalized StationXML format (FDSN, 2019). Today, all A-nodes produce StationXML metadata and most of Résif-SI are moving toward a common toolbox and database for metadata edition. The Metadata editing made easy: YASMINE section gives additional information on this project, which has high priority in Résif-SI, because StationXML gives visibility and full acknowledgment of all dataset contributors—from station to data distribution (see the Data fairness section). This normalized information is automatically exported to the Résif-SI web portal and to other services. Moreover, StationXML allows referencing to persistent network identifiers to facilitate citation and includes new fields necessary, for example, for OBS data (see the Marine data section).

Figure 4. Technical architecture of Résif-SI. A-nodes collect and validate the raw data, manage real-time data flow, edit the metadata, and submit them to the B-node Résif-DC. Résif-DC concentrates the real-time data flow, which is used by the national seismic alert system operated by the Commissariat à l'énergie atomique et aux énergies alternatives (CEA) (on top of their own dedicated data flow for the networks they manage), stores data and metadata in the long term, distributes it through standardized and specific media, develops side products. Federated into EIDA, open data and metadata are accessible by end users through a federator. The color version of this figure is available only in the electronic edition.

Data collection and completion

The management of data flows is specific to each A-node. However, Résif-SI offers its own tools that are generic enough to be deployed outside their initial context. An example is MOdular aRchive bUilder from Multiple Origin Temporal Traces & Other stuff (MORUMOTTO) (Geber, 2019), developed by IPGP for A-node Volcano. MORUMOTTO is used by network operators to quality control a data archive, and, in particular, to detect and correct data gaps and overlaps. Data are regularly fetched from a pool of different sources.

Data validation

Data validation procedures are specific to each A-node.

Data from networks FR and RA are qualified "M" ("data center modified, time-series values have not been changed," see Ahern *et al.*, 2012, p. 108), and is made available three to five days after collection by A-nodes, after having undergone the following checks:

- data have been completed as much as possible;
- there is no overlap, even at the sample level;
- instrument response validity epochs do not overlap;
- the PSD of the signal is in accordance with what is expected for each particular station.

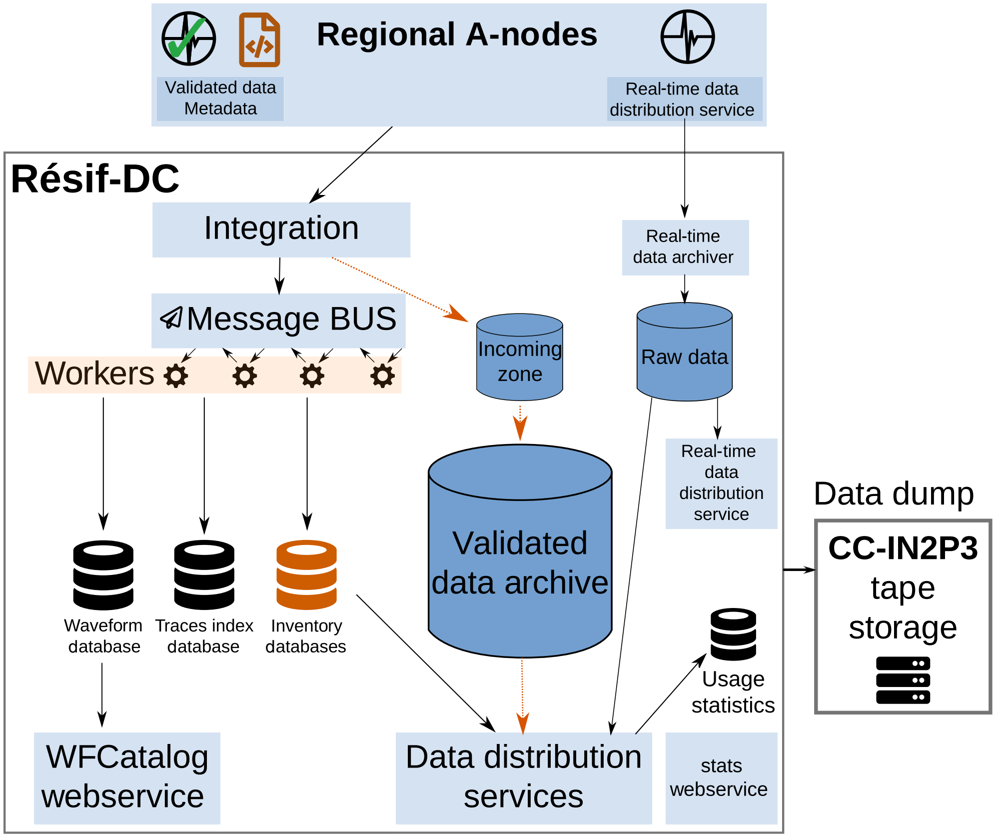


Figure 5. Validated data and metadata are integrated at Résif-DC, by passing through a series of workers, which ensure that the data and metadata conforms to the specifications, register it in the archive and in the database, index it, and create the side products. Raw data are concentrated, registered, and made available through standardized webservices with very low latency (around 10 s). Data, metadata, and other assets are dumped on a distant tape system at the Computing Centre of the National Institute of Nuclear Physics and Particle Physics (CC-IN2P3, Lyon), to secure them from a local disaster. The color version of this figure is available only in the electronic edition.

Data from networks G, GL, MQ, WI, and PF are made available with the quality label "Q" ("quality controlled data, some processes have been applied to the data;" see Ahern *et al.*, 2012, p. 108), and undergoes additional checks, namely:

- waveform modeling of teleseismic earthquakes of magnitude larger than 6 through the Seismic source ChAracteristics Retrieved from DEConvolvolving teleseismic body waves (SCARDEC) method (Vallée *et al.*, 2011) shows a good agreement between the observed and predicted signals;
- for multiparametric stations, the acceleration-converted seismometer and accelerometer waveforms are identical for a selection of suitable earthquakes.

Because of the complexity of these additional operations, validated data from these networks are available 6–12 months after its collection.

Availability of validated data from SisMob temporary networks depends on the specificity of each experiment, and generally varies between 1 and 6 months (SisMob data might be under embargo for up to three years).

Table 2 summarizes the latency for the production of validated data by A-nodes and their data quality label. For networks collected in real time (see the Real time: SeedLink section), raw data are immediately available and eventually replaced with validated data.

Quality of service

The Résif-SI organization does its best to implement a set of high-quality services with high-operational performance, that is, as continuously available and robust as possible. To achieve this, we developed a culture of quality of service, based on a set of the best practices, among which:

- a single help desk, accessible by email at resif-dc@univ-grenoble-alpes.fr. Users are typically scientists searching for data, data or metadata producers needing help to access the services, or network managers needing assistance.
- a clear presentation and documentation of our services through the website (see Data and Resources);
- an automated system that tests thoroughly our services and infrastructure, and monitors the overall activity to alert the relevant personnel of Résif-SI, as early as possible, in case potential issues or anomalies are detected;
- a Really Simple Syndication (RSS) feed to publish news or scheduled downtime events.

The website (see Data and Resources) is also the place to browse all the available metadata, to get information about citations, data, and metadata access. It is also used for announcements regarding operational news of Résif-DC, such as maintenance downtimes or new services availability.

The monitoring system checks all the Résif-DC services, mimics user access for testing, and alerts the operational team, if necessary. It also monitors lower levels of the infrastructure of Résif-DC. The gathered information is used in reports, to publish the overall availability of the data center.

TABLE 2
Latency for the Production of Validated Data by A-Nodes

A-Nodes	Data Quality	Latency
RLBP	M	3–5 days
RAP	M	3–5 days; 1 yr for the data from stations in buildings and borehole
GEOSCOPE	Q	6–12 months
Volcano	Q	6–12 months
SisMob	M	1–6 months

TABLE 3
Average Service Availability for Year 2020 Based on Network Reachability

Service	Availability (%)
fdsnws-dataselect	99.84
fdsnws-station	99.62
fdsnws-availability	99.60
Real-time data	99.87
resifws-timeseries	99.66
resifws-timeseriesplot	99.65
eidaws-wfcatalog	99.29

Availability of 0.1% corresponds to 9 hr.

Measuring and publishing the availability of our services is an ongoing project that will be completed in 2021. Service level agreements will be defined by the executive committee for our main services, and the data center's yearly report will show the metrics. Table 3 provides a rough estimate of service availability at Résif-DC in 2020, based on network reachability. Based on feedback from the EIDA User Advisory Group, EIDA is presently working on a periodic test bench of all the nodes based on random requests of data and metadata, which will be more representative of the services availability.

Data Management: Résif-DC

The construction of Résif-SI is patterned on a data management plan (e.g., Michener, 2015) describing the responsibilities of each partner, how long data should be stored, how metadata are managed, how data are accessed. Résif-DC has a central role to play in this plan: It ensures that data are securely stored and easily accessible.

Data archive description

Résif-DC stores the data in miniSEED (using the SDS structure, Ahern *et al.*, 2012; GFZ and gempa GmbH, 2008) or PH5 format (Hess *et al.*, 2018). The latter is designed by IRIS Portable Array Seismic Studies of the Continental Lithosphere (PASSCAL), which is based on HDF5 (Folk *et al.*, 2011) and widely used by the temporary experiments using dense arrays.

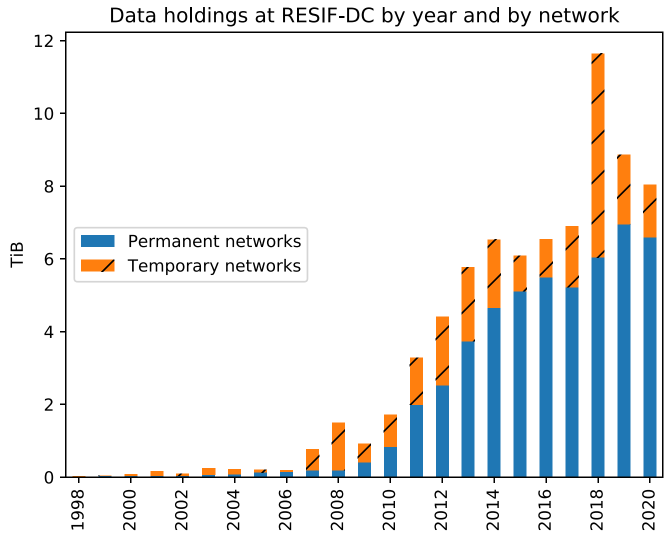


Figure 6. Current validated data holdings at Résif-DC grouped by year. This figure illustrates the growth of the data produced by seismological networks since 1998. The color version of this figure is available only in the electronic edition.

Résif-DC holds miniSEED data from 1986, all accessible through the FDSN dataselect webservices. Figure 6 shows the volume of miniSEED and PH5 data, by network, held for each year as of January 2021. We account for 68% of data from permanent networks versus 32% from temporary experiments.

The increasing amount of data stored came from two major factors: first, the deployment of, approximately, 150 new permanent broadband stations; second, the increasing number of temporary experiments associated with SisMob. We anticipate a strong data growth in coming years, due to increasing number of experiments using dense arrays and the foreseen integration of data from distributed acoustic sensors, but, the numbers are still difficult to quantify, as discussed in Quinteros *et al.* (2021).

The data management description has been used to design an efficient central storage at Résif-DC. Data are kept on two classes of storage, ensuring the same data security but with different levels of performance. Although, seldom-requested data are still quickly accessible, they are kept on a less powerful and less expensive media. This storage tiering is one of the many features of our storage provider Stockage Unifié Mutualisé Massif Evolutif et Réparti (SUMMER) and is totally transparent for the data center systems, thanks to a storage abstraction layer (autofs). Once a year, the data that can be transferred to less powerful storage media are selected according to the seismic network's data management plan and manually moved.

Securing the data

When it comes to store the data securely, one has to identify and address several types of risks. Résif-DC is concerned about:

1. hardware failures;
2. unintentional or malicious alteration or deletion of the data by an operator;
3. local disaster destroying physically the archive located in Grenoble (flood, earthquake, etc.).

Risks 1 and 2 are addressed de facto by Résif-DC through the choice of storage provider SUMMER (Université Grenoble Alpes, 2016), operated by the University of Grenoble Alpes, which secures the data in three different data centers and offers a daily snapshot service. Using the university storage service, we also ensure that all the data are kept on an academic and publicly funded platform.

The risk 3 of a local destruction has been addressed by copying twice a year a snapshot of our archive to tapes stored at a distant place. At this juncture, we have secured a collaboration with the Computing Centre of the National Institute of Nuclear Physics and Particle Physics (CC-IN2P3, IN2P3/CNRS, 1976) located in Lyon, to serve as this remote site. Although, the recovery plan is not fully formalized, we guarantee that the data hosted at Résif-DC will survive the mentioned risks, as long as it has been stored for, at least, six months. An additional copy for recent data are ensured by A-nodes, which are committed to keeping a secure copy of their datasets for a duration of two years.

Altering or destroying the data hosted by Résif-DC would require an operator or an attacker to get access to high privileges on three independent infrastructures, namely:

- Write access to the Résif-DC infrastructure would allow to delete live data in the archive. The data can in that case be retrieved from SUMMER snapshots or backups.
- Write access to the SUMMER infrastructure would allow to destroy backups, snapshots, and live data. The data and metadata can, in that case, be recovered from a distant copy made twice a year at CC-IN2P3.
- Write access to the CC-IN2P3 infrastructure would allow to destroy the distant copy.

Résif-SI considers that the scenario of the combined access to these three infrastructures is unlikely to happen.

Data distribution services

Résif-SI data distribution is based on a small set of elementary and very robust data services, which can be divided into: real-time data access, asynchronous data access, and webservices. These base services constitute the building blocks for higher-level data services and products provided by Résif-SI or by European and international partners. Résif-SI is fully integrated into EIDA, through its elementary data services. EIDA offers a wide set of additional services and tools, including smart clients, federated archive access, and web interfaces. A full description of EIDA can be found in Strollo *et al.* (2021).

The Résif-SI web portal (see Data and Resources) provides information about Résif-SI, documentation on how to use the elementary and additional services, licensing, and so forth. It also provides URL builders for the webservices and dynamic search options for browsing the seismic networks, station metadata, data availability, data quality, and so forth.

Résif-DC has implemented most of its data models and workers to preserve the specificity and accuracy of the metadata produced by A-nodes. In particular, care was taken to be able to export the rich Résif-SI metadata. This a priori expensive solution proved to be profitable in the long term, because it enabled Résif-SI to integrate from early on the description of nonseismological channels (polynomial responses), and to manage all station and channel comments, as well as OBS metadata. In this way, additional services are effective across all the data types managed by Résif-SI.

The two initial FDSN webservices station and dataselect (FDSN, 2013) were put into production in 2013. Their implementation is based on IRIS WebServiceShell (Incorporated Research Institutions for Seismology [IRIS], 2016), which we interfaced (via Python) with our databases and archives. All the other webservices have been implemented in Python and ObsPy (Beyreuther *et al.*, 2010; Megies *et al.*, 2011; Krischer *et al.*, 2015) within the Flask framework (Ronacher, 2010), and their codes are fully available on public code repositories (see Data and Resources).

The basic Résif-DC data services are:

- A SeedLink server (GFZ and gempa GmbH, 2008) (standard port 18000), for real-time access to miniSEED data (see Data and Resources). This service, based on IRIS ringserver (Trabant, 2011), provides a single access point to all the real-time data streams of the A-nodes for which metadata are publicly available.
- The standard fdsn-dataselect webservice (FDSN, 2013) to access validated or near real-time data—public or restricted—for all of the miniSEED archive.
- The FDSN-compliant ph5-dataselect webservice (Résif, 2019a), which serves large-*N* data (stored in PH5 archives) as miniSEED or Seismic Analysis Code (SAC) format on the fly (useful for small subsets of the PH5 datasets).
- The standard fdsn-station webservice (FDSN, 2013), to retrieve station metadata.
- The standard fdsn-availability webservice (FDSN, 2013), to interrogate the miniSEED data inventory.
- The FDSN-compliant ph5-availability webservice (Résif, 2019a), to interrogate the PH5 data inventory.
- The eidaws-wfcatalog webservice (Trani *et al.*, 2017), to retrieve data availability as well as other metrics (e.g., gaps, overlaps, root mean square [rms]);

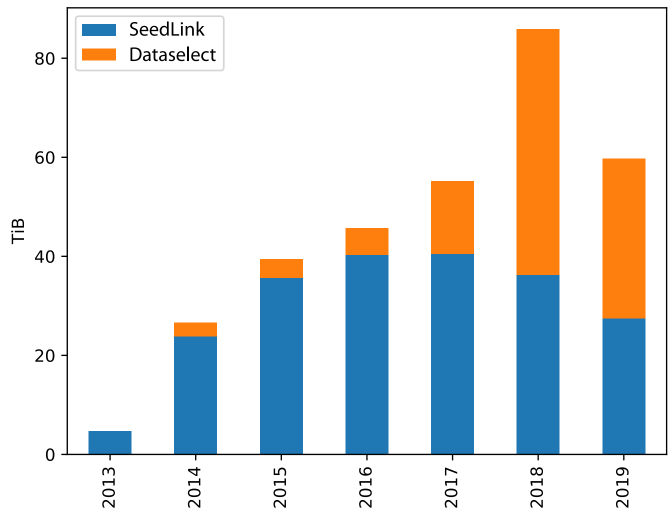


Figure 7. Amount of waveform data distributed yearly through SeedLink (real time) and the International Federation of Digital Seismograph Networks (FDSN) dataselect. Note that statistics from the now retired EIDA Arclink services are not included. The color version of this figure is available only in the electronic edition.

- The standard fdsn-event webservice (FDSN, 2013), operated by Bureau Central Sismologique Français, French Central Seismology Bureau – Réseau National de Surveillance Sismique, National Seismic Monitoring Network (BCSF-RéNaSS), which gives access to event parameters located by BCSF-RéNaSS (Masson *et al.*, 2021).

Résif-DC also allows the retrieval of a complete dataset, by network, using rsync protocol. The user is granted read-only access to the part of the archive exposing the dataset. This is, particularly, useful to retrieve entire datasets in PH5 format. Figure 7 shows the amount of waveform data served by fdsn-dataselect and SeedLink.

Built on top of these base services, the following additional webservices are available from Résif-DC:

- The time series and time-series plot webservices (Résif, 2019a), to obtain preprocessed waveforms or plots.
- The resp, sacpz, and evalresp webservices (Résif, 2019a), to obtain or plot instrumental responses.

A complete catalog of the Résif-DC webservices is available (see Data and Resources). More details and usage examples are given in the Data Access Details and Examples section.

Data fairness

Data fairness refers to the Findable, Accessible, Interoperable, Reusable (FAIR) guiding principles for scientific data management and stewardship (Wilkinson *et al.*, 2016). We here discuss how far we are complying to those best practices.

The findability of our data and metadata is ensured by the international standard station webservice (FDSN, 2013), accessible via any web browser, http client, and the community's most popular tools like ObsPy (Beyreuther *et al.*, 2010; Megies *et al.*, 2011; Krischer *et al.*, 2015), the EIDA's fdsnws_scripts (Heinloo, 2018), or the webservice fetch scripts (Hutko, 2013). In addition, most of the Résif datasets have digital object identifiers (DOIs) associated with their seismic network, according to the standard procedure approved in 2014 by the FDSN (Evans *et al.*, 2015). Within Résif-SI, much effort has been dedicated to manage the quality of the published metadata, as stated in the Metadata enhancement section. The seismological metadata are very descriptive, allowing scientists to identify precisely the data needed. In addition, the station webservice provides a rich set of options to make selections in the metadata catalog, such as time period, geographical region, or update time. Findability of Résif-SI data is also guaranteed by the ORFEUS European metadata catalog (ORFEUS, 2020a), implemented by EIDA and maintained by ETH-Zürich, which enables users to fetch metadata from all the EIDA's data centers in one request, at one entry point. Finally, within FDSN, a large effort in establishing data routing tables has been achieved. Now, each seismological network has a well-known reference data center, and this information is made publicly available by each data center and can be consulted for each network on the FDSN's network details page (FDSN, 2020).

The accessibility of the data is ensured by the international standard dataselect webservice (FDSN, 2013) that gives access to all Résif's archive, except for data in PH5 format, which has its own nonstandard ph5-dataselect webservice (Résif, 2019a). Authentication is also provided for restricted data, with local methods or with EIDA's authentication system.

The interoperability is ensured inside Résif, by adopting a common vocabulary for StationXML metadata, which makes it possible to expose on the Résif portal significant parts of the StationXML content coherently across all the operators. For the services to the user, interoperability is first and foremost ensured by webservices in strict adherence to FDSN standards when available. This is the key for universal data and metadata access through interoperable additional layers across many data centers offering the same services. At a service level, EIDA offers such interoperable services (e.g., WebDC3, ORFEUS 2020b, which gives transparent access to the data of all the EIDA nodes), whereas, many users develop workflows based on the FDSN webservices interfaces (e.g., Zaccarelli *et al.*, 2019; MacCarthy *et al.*, 2020). Cross-disciplinary interoperability requires challenging standardization of metadata vocabulary, both for data and services. The interoperability between the different Résif data types (seismology, GNSS, gravimetry) is delegated to the EPOS infrastructure, into which Résif is fully integrated. National technical discussions are undertaken across different EPOS activities in France, rather than strictly within Résif perimeter, and the Résif seismology and GNSS community actively contribute to EPOS. The EPOS cross-disciplinary interoperability is achieved through a dedicated layer that connects different services through a metadata catalog that uses standardized vocabulary to describe the services. In addition, the scientific users of Résif data are already observed to mix data and data products from different domains (e.g., seismic waveforms with environmental data, such as weather condition or ocean-wave activity) within smart clients that they create, or by mixing data from local files downloaded from different sources with direct download of waveform data. The main condition for the success of these applications is meticulous application of domain standards, and the effective data and metadata accessibility.

The data provided by Résif-SI are also made reusable in respect of the FAIR guidelines. The rule of Résif-SI is to distribute all open data under the Creative Common Attribution 4.0 license (Creative Commons, 2013) CC4.0:BY, coherent with French law. StationXML specification (currently at v.1.1; FDSN, 2019) does not provide a field to indicate the data license, but we made this information easily accessible, by systematically exposing the DOI of a network via the <Identifier> tag in StationXML metadata. As an example, metadata from RAP network (network code "RA," see Data and Resources) contains the following tags:

```
<Description>RESIF-RAP Accelerometric permanent network</Description>
<Identifier type="DOI">10.15778/RESIF.RA</Identifier>
```

License information is provided in the DataCite XML document (DataCite, 2019) of each network via the <rightsList> tag. For instance, the "RA" network DataCite XML (see Data and Resources) contains the following tags:

```
<rightsList>
<rights rightsURI="info:eu-repo/semantics/openAccess">Open
Access</rights>
<rights rightsURI="https://creativecommons.org/licenses/by/4.0">
Creative Commons By 4.0 Universal</rights>
</rightsList>
```

Citation instructions are available on the DataCite page (e.g., see Data and Resources) and dynamically presented by several portals (e.g., FDSN network details, FDSN 2020, or Résif seismic data portal, re3data.org, 2016). Résif-DC provides a list of citation instructions in the portal's citation page (Résif-SI, 2020). We strongly encourage scientists to cite data producers and distributors in their papers.

The use of licenses for seismic waveform data is presently only at its beginning, with licenses being put in place for many seismic networks in Europe. A strong motivation is not only the citations (see earlier), but also to handle liability issues in the case of erroneous data or metadata. Résif-SI is actively working, within the framework of a cooperation between ORFEUS and IRIS, to promote good citation usage to scientific users of waveform data, and to inform publishers of these progresses, as the introduction of licenses engages the liability of the journals and of the publishing scientists, even though this liability is not yet reinforced.

An accompanying issue is that data centers that hold copies of the datasets need to expose the license information in a proper way. For this to be practically possible, Résif-SI has, with the agreement of the involved organizations, stopped all data copies to other data centers, which were previously put in place, and asked for the old copies to be deleted. The only exception is GEOSCOPE data archived at the IRIS Data Management Center (DMC), due to a strong user base of GEOSCOPE data through this data center. Overall, the seismological community worldwide still needs to efficiently communicate license information to users, but, if the citations (see earlier) are properly done, the license CC4.0:BY is respected.

Citation through the use of DOIs is only meaningful if the associated metadata are sufficiently rich to include all the parties that contribute to the data production, management, and distribution. Résif-SI uses the contributor field of DataCite metadata (DataCite, 2019) to acknowledge different types of contribution and roles within each seismic network, and, more specifically, the contributorTypes DataManager (Résif-SI), Distributor (Résif Data Center), HostingInstution (Université Grenoble Alpes), DataCurator (the relevant A-node), DataCollector (organizations that operate the seismic network), and Sponsor (funding sources).

From logging to usage statistics

Data usage statistics are important information for the different persons associated with data production, management, and distribution. For example, the Résif-DC team needs to have real-time usage statistics, correlated with IT system metrics to catch anomalies or anticipate requests growth, whereas, data producers, project principal investigators, and funding agencies need compiled usage statistics in a larger view and for larger time span, such as overall statistics (e.g., number of requests, users, shipped volumes, countries of request) or network relevant information (e.g., network level download statistics, the most accessed stations, countries of requests).

To satisfy all the needs for statistics, we built a system that concentrates the information, builds statistics, and gives access to dashboards (Fig. 8).

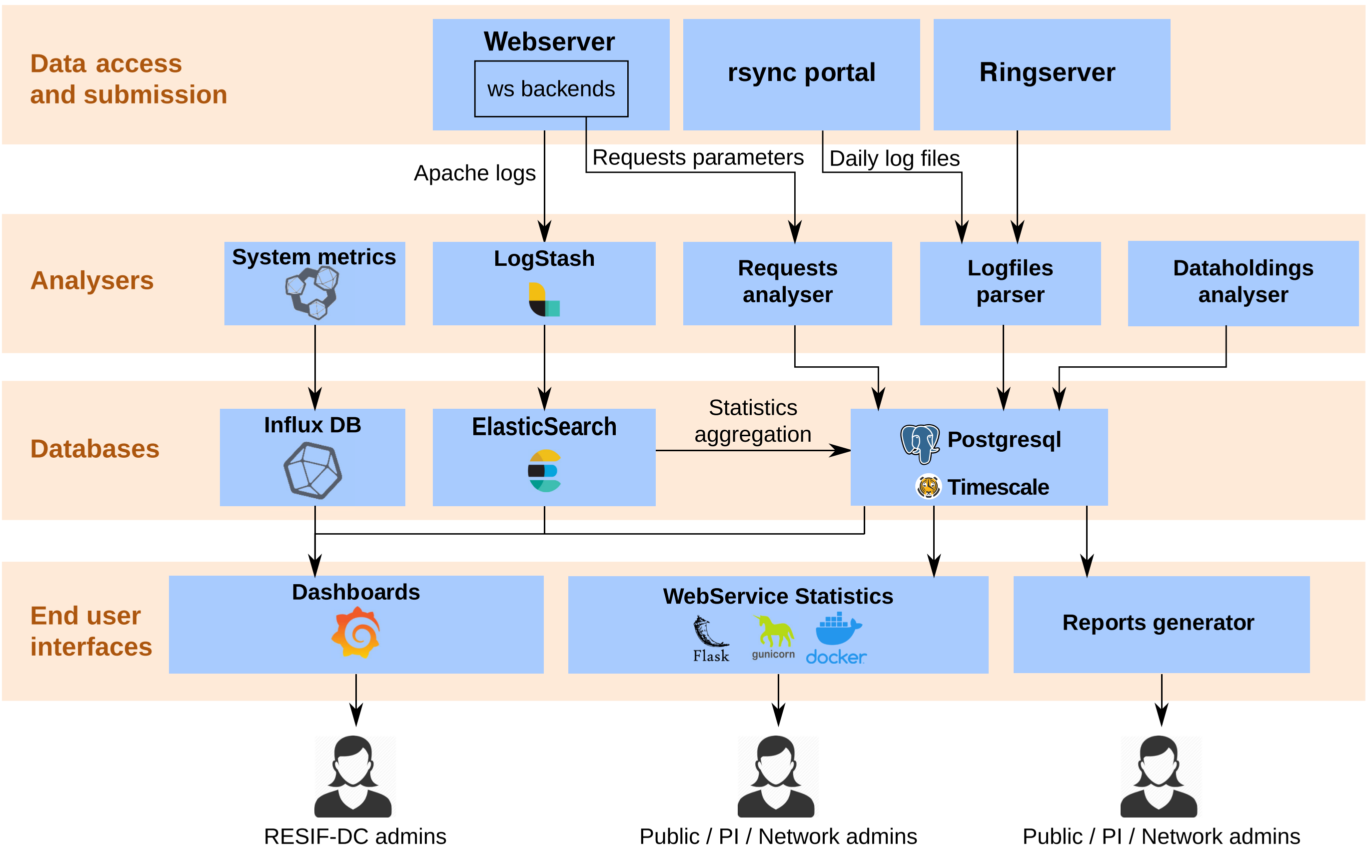


Figure 8. The logging system analyzes events from heterogeneous sources (e.g., web access logs, webservices requests, and server logs); the statistics are computed by different analyzers and aggregated in one central database (Postgresql). The statistics can be queried by various interfaces intended for different audiences; in particular, a public webservice resifws-statistics is available. The color version of this figure is available only in the electronic edition.

The constraints come from the variety of data access methods, each one having their own log formats that are not always compatible with real-time processing. Another difficulty is to analyze the data served through webservices. It is easy to capture the quantity of data shipped by the web server, but it is not possible to know what it is made of (which network, station, location, or channel). Consequently, we have to analyze each request and evaluate the typology of the response. All the gathered information is stored in a PostgreSQL database, aggregated, and anonymized. The anonymization concerns the client's IP address and consists of hashing it as soon as it enters the databases. Then, when the requests are aggregated, we only keep HyperLogLog objects (Flajolet *et al.*, 2007) to compute the cardinality of the clients. Therefore, our databases respect the European legislation (General Data Protection Regulation), because there is no way to retrieve an IP address or any kind of personal data.

The statistics can be accessed through several means: we provide a webservice for end users (see the Data Access Details and Examples section) and interactive dashboards for internal usage (e.g., analyzing operational events in real time, usage, and performance evaluation, Fig. 9). Figure 10 is an example of rendering statistics in a geographical map. Recently, we published a system allowing automatic creation of plots and usage statistics in form of a report (resif-delivery-stats-plotter, Bollard, 2021). This program is aimed at network managers or principal investigators seeking information on data and metadata usage.

New Developments

Large-*N* data management

Large-*N* data are produced by temporary experiments deploying small, self-contained seismometers (the so called "nodes") in dense arrays, with high sampling rate (i.e., above 100 samples per second, e.g., Brenguier *et al.*, 2016; Dougherty *et al.*, 2019; Gimbert *et al.*, 2020). The raw data generated by those experiments is, approximately, 1 GB per day per node, and an experiment can use hundreds of nodes for several months. In addition, the data management needs specific workflows and data format treatments.

To host and distribute this new kind of data, the SisMob A-node developed a specific workflow that produces validated data in PH5 (Hess *et al.*, 2018)—a data format developed at IRIS by PASSCAL and commonly used for large-*N* data. To ingest this new data format, Résif-DC also adapted the integration mechanisms and storage repositories, because PH5 data cannot be indexed nor referenced in the same way as

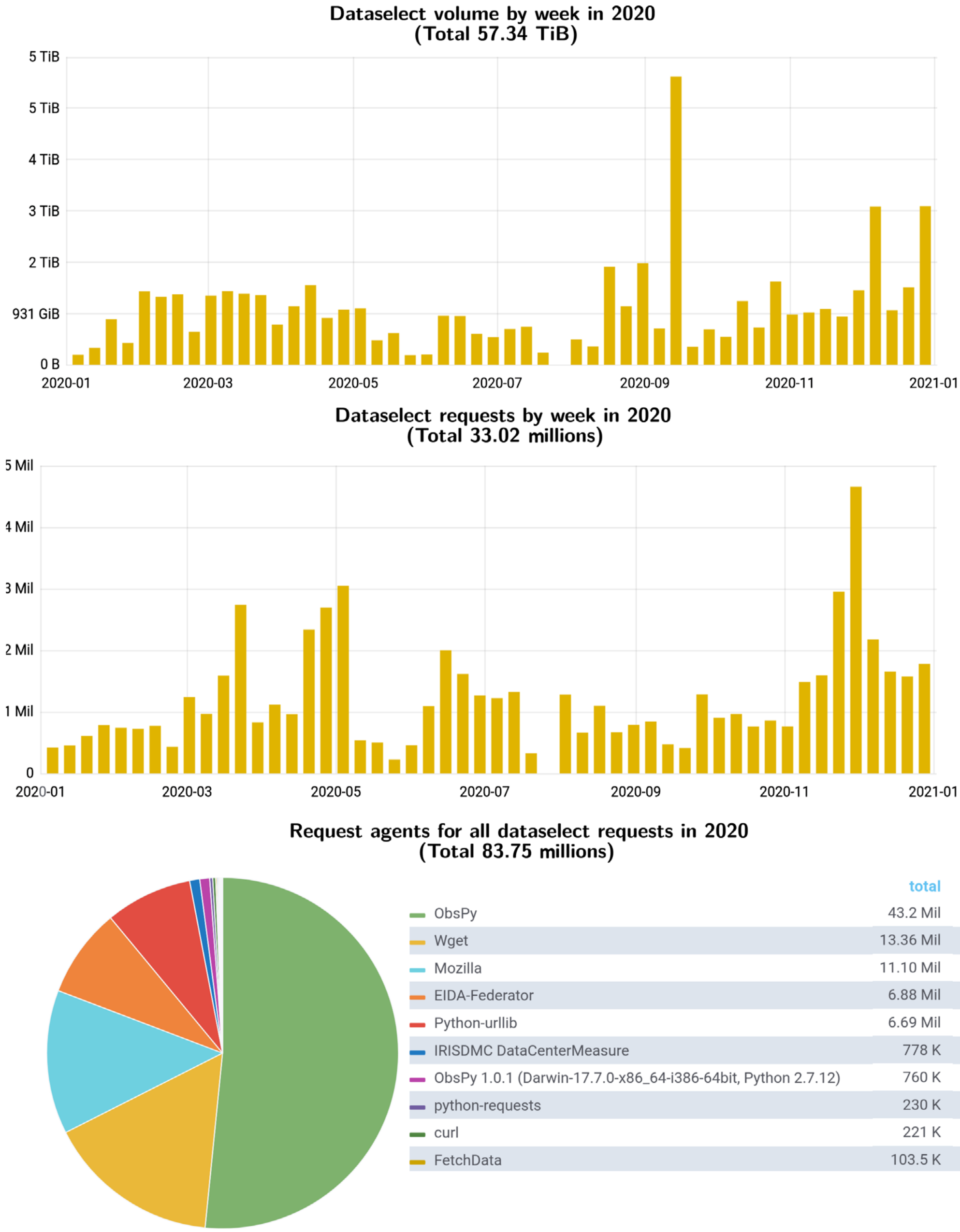


Figure 9. This dashboard view is an example of the web interface presenting some of the Résif-DC distribution metrics. Here, the figures focuses on the fdsnws-dataselect webservice during 2020, showing the volume sent (51.34 TB), the number of successful requests (33.02 millions), and the distribution of user agent requests for all the requests (83.75 millions). This last number accounts for the successful requests to the /query and /queryauth methods (HTTP 200), the requests returning no data (HTTP 204 and 404), the requests to other methods (documentation, /auth, wadl file), the requests producing errors due to improper parameters (HTTP 400 and 401) or insufficient permissions (HTTP 403). The color version of this figure is available only in the electronic edition.

miniSEED data. Furthermore, to host the foreseen volumes, a dedicated archive storage has been setup.

The metadata, on the other hand, is created in classical StationXML format, as other miniSEED data, and submitted at Résif-DC using the standard procedures described previously.

The PH5 archive is served to end users by two means:

- rsync access for a restricted set of users needing to download the entire dataset; and
- the ph5-dataselect webservice (Résif, 2019c) allowing the selection of a subset of the data in miniSEED format with the same options and syntax as the FDSN dataselect webservice specification.

Résif-DC also provides the ph5-availability webservice (Résif, 2019b) for the users to get the time span of the available data.

Marine data

Marine data collected on autonomous sensors, such as OBSs, follow specific guidelines developed and distributed through the FDSN "mobile instrumentation" working group and the European Union projects EPOS/ORFEUS, Environmental Research Infrastructures (ENVRI)-FAIR, and Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA). These guidelines include: data quality labels indicating whether the data were corrected or not for the instrument clock drift; standards for postimplementing leap seconds; component code standards for horizontal channels that are not geographically oriented; orientation and dip standards for pressure channels; and station naming rules for repeated deployments at the same station (see Clinton *et al.*, 2018, their Appendix B, for the latest published version and Crawford, 2019; for the latest proposed version). These guidelines have been adopted already by some OBS parks and EIDA data centers. The AlpArray OBS component has been archived following these guidelines.

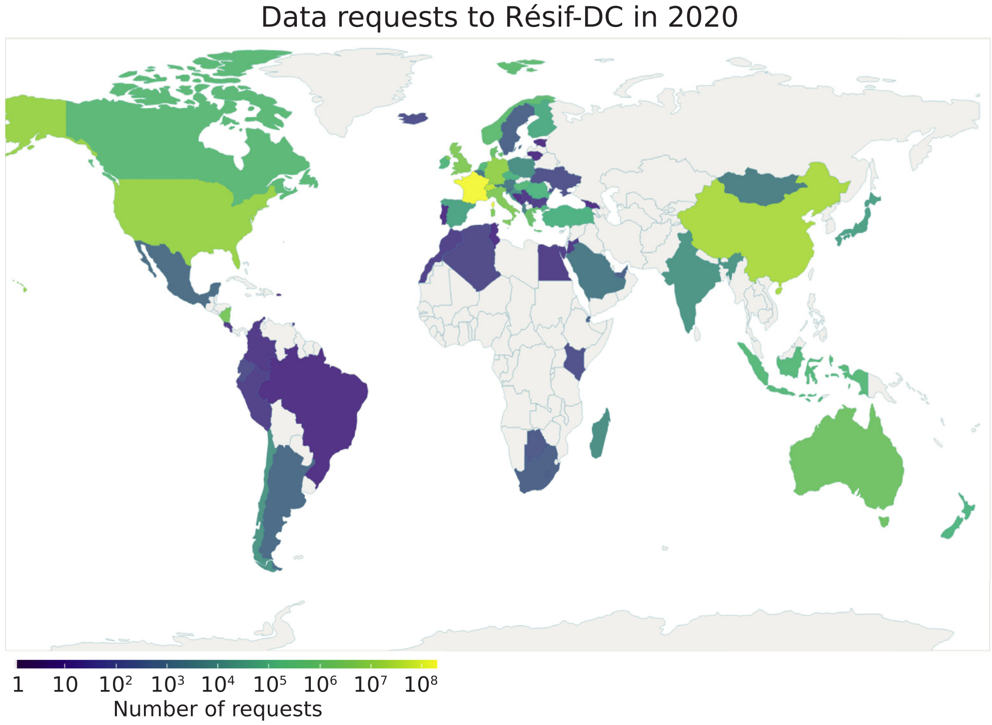


Figure 10. Geographical distribution of data requests to Résif-DC in 2020. The color version of this figure is available only in the electronic edition.

Résif has been working, since 2017, on the integration of data from French OBS parks. This integration will be accomplished very soon by the commissioning of the dedicated MARINE A-node. Within this frame, a system for creating FDSN-standard data and metadata for OBSs using standardized, easy-to-read information files is currently in development (Crawford *et al.*, 2019). A specific visual quality control is also in development (Goubier and Crawford, 2021), to allow the instrument providers and scientific users to verify instrument responses, noise levels, and time corrections.

To date, Résif-DC distributes data from two temporary campaigns:

- Réunion Hotspot and Upper Mantle – Réunions Unterer Mantel (RHUM-RUM) (Barruol *et al.*, 2017): 57 OBS for two months of data, three components at 50 Hz
- AlpArray (AlpArray Seismic Network, 2015): Eight OBS for eight months of data, three components at 50 Hz.

Two other datasets are in the process of validation and integration:

- Sismantilles (Laigle *et al.*, 2007): 20 OBS for four months, three components at 50 Hz.
- European Multidisciplinary Subsea Observatory - Monitoring the Mid-Atlantic Ridge (EMSO-MoMAR) (IPGP, 2007): Five OBS for 10 yr three components at 50 Hz.

OBSs cannot obtain precise Global Positioning System (GPS) timing during their deployment. The instruments are equipped with a very accurate clock to minimize the problem; however, this clock drifts on the order of 1–2 s/yr (e.g., Loviknes *et al.*, 2020). The standard protocol for OBS time correction is to synchronize an OBS with a GPS signal immediately before deployment and after recovery. The measured timing deviation is assumed to have accumulated linearly over the deployment interval; therefore, the applied correction for time drift is linear. This assumption has been checked by Hable *et al.* (2018) for the instruments used during the RHUM-RUM experiment (French and German pools), using ambient noise correlations.

Some users, however, prefer to have the data "unmodified," even if this means that there are timing errors of the order of a second. That is why Résif-DC delivers both types of data, giving quality label "Q" (quality controlled data, some processes have been applied to the data; see Ahern *et al.*, 2012, p. 108) to those that are corrected and "D" (the state of quality control of the data is indeterminate; see Ahern *et al.*, 2012, p. 108) to those that are not (see the Specificity of OBS data and metadata section).

Organizing and improving archival of data from building and infrastructure monitoring

Since 2010, Résif has been involved in specifying metadata and solutions for building and infrastructure monitoring (Clinton *et al.*, 2018). More and more high-quality seismic sensors are deployed across structures, with the same technical issues than for classical seismological networks (in large numbers, with continuous recordings) and using, therefore, seismological standards. The StationXML metadata specification (FDSN, 2019) can be used to include specific information on the structure, allowing the engineering seismology community to integrate infrastructure monitoring data into standardized practices. The building description, which is key for engineering purposes, can be included at two levels: (1) following the European Macroseismic Scale typology of buildings, based on the material of construction (Grünthal, 1998), and (2) through a full description of the building characteristics according to the Global Earthquake Model (GEM) Building taxonomy (Brzev *et al.*, 2013).

To date, Résif-DC distributes seismic data from five instrumented buildings:

- City-hall of Grenoble, France (Michel *et al.*, 2009; Guéguen *et al.*, 2020): Six three-component (3C) accelerometric sensors since 2004 (stations OGH1–OGH6), one additional 3C sensor at intermediate height (station OGH7), and one weather station at the top since 2019 (station OGH8).
- Ophite Tower in Lourdes, France (Michel and Guéguen, 2018): Three 3C accelerometric sensors and 15 1C accelerometric sensors distributed over the building height since 2008 (station PYTO) plus one temperature sensor at the top.
- Prefecture building in Nice, France (Lorenzo *et al.*, 2018): Two 3C accelerometric sensors and 18 1C accelerometric sensors distributed over the building height since 2010 (station NCAD), including one free-field station.
- Basse-Pointe College in Martinique Island: Two 3C accelerometric sensors and 18 1C accelerometric sensors distributed over the building height since 2010 (station CGBP), including one temperature sensor and one free-field station.
- Centre de découverte de la Terre in Martinique Island (Gueguen, 2012): Two 3C sensors in trigger mode in a specific building with rubber bearing since 2005 (stations CGCP and CGLR).

Metadata editing made easy: YASMINE

For almost 20 yr, the IRIS PDCC (Casey, 2016) was the main standalone graphical user interface (GUI) tool available to create and maintain station metadata in SEED dataless format (Ahern *et al.*, 2012). In 2017, several years after the adoption of the FDSN new StationXML metadata standard (FDSN, 2019), there was still no solution widely available to edit and create native StationXML files. This slowed down the adoption of StationXML by, at least, all Résif-SI contributing A-nodes and forced us to use an interim SEED dataless to StationXML conversion solution. This problem was hindering Résif-DC being able to export rich metadata to the user, including, for example, the DOI of the network, or the contributing organizations.

Résif-SI first collected the metadata creation and edition needs among the French community and closest partners. The requirements were then generalized so that they fulfill not only Résif-SI needs but also a much broader international community.

In 2018, IRIS-DMC contracted Instrumental Software Technologies, Inc. (ISTI), to build a tool for StationXML creation and editing, and this first software was already satisfying some of Résif-SI requirements. Résif-SI then contracted ISTI to continue the development of Yet Another Station Metadata INformation Editor (YASMINE), to satisfy our requirements. The results are two independent pieces of software: yasmine-GUI and yasmine-CLI.

One of the main innovations was the introduction of the new atomic response object library (AROL, Wolyniec *et al.*, 2019). This library, written in YAML Ain't Markup Language (YAML), is contributed by Résif-SI and is the conversion of the long-maintained PZ format library. Each stage of a device is only defined once and then linked to the definition of the many different instrument configurations possibilities.

The web-based GUI yasmine-GUI offers the user the ability to create and edit StationXML metadata. The user can create files from scratch through a guided wizard process and can import response files from either nominal responses for seismic instruments library (Templeton, 2017) or AROL. The user can also modify any part of the instrument response. A comparison mode allows the user to compare the instrument response of two StationXML files, sharing the same station-channel-network-location. Often used description elements (vault, geology, and comments) can be stored in the General ATomic lIbrary of Tiny Objects (in YAML, Saurel *et al.*, 2019), to simplify the task of homogenization of those elements over several StationXML files. Finally, the user can store any of the network, station, or channel elements in a user library that will store them as templates for further reuse.

The command-line based yasmine-CLI allows the user to modify an existing StationXML file. The user can add, delete, or modify StationXML elements, except for the instrument response elements. The user can also split a complex StationXML into multiple simpler files (e.g., split a network StationXML into several perstation files) and print the instrument responses of each channel contained in the StationXML. This tool will allow many automated StationXML file modifications from scripts.

The two pieces of software are written in Python and rely heavily on ObsPy for the StationXML content manipulation. They are both compatible with the latest StationXML 1.1 standard and are released under the GNU General Public License (GNU GPL) version 3 and distributed as Python packages. The Résif-SI installation of YASMINE is reachable, see Data and Resources.

Conclusion and Future Challenges

The choice of a distributed yet strong architecture for the national Résif information system in seismology, with clearly identified roles and responsibilities, has proven effective for France, to deliver seismic waveform data to users worldwide. The service to users has significantly improved in terms of features offered and service robustness with minimal changes of running costs and keeping the same number of permanent staff members. The advantage of the system is that it capitalizes on the distributed human resources and competence nationwide, and maintains visibility for each involved institution. The cost of this architecture is the continuous efforts needed to maintain the technical coherency and cooperation across all institutions, and to overcome communication difficulties across the distributed system. The choice made very early on to preserve original data models allowed us, despite the initial cost, to respect the specific constraints on data and metadata for all the French networks—an important consideration for the cohesion of Résif-SI and for the scientific users. Based on this experience, the Résif board of directors has recently chosen a similar architecture for information on earthquakes, with identified roles of different operators responsible for warning, bulletins, catalogs, and shake maps, rather than each institution creating their own (Masson *et al.*, 2021).

Résif faces future challenges, related to two main issues. First, the success of the system at a national level means that the quantity and complexity of collected and distributed data is ever-increasing. For example, all broadband OBS data will, in the future, be integrated into Résif. The second issue is an order of magnitude change in data integration associated with large-N and distributed acoustic sensing equipment. These challenges are international, and Résif will continue to engage in international discussions on how to tackle them (see Quinteros *et al.*, 2021). We believe that the Résif architecture is sufficiently solid to accommodate the changes through engagement of all the partners. The experience from Résif shows that a distributed but strongly organized information system is a good architecture for countries that face the challenge of integrating and distributing data across many organizations.

Data and Resources

All the software produced by Résif-SI is open-source and available at https://gitlab.com/resif and https://gricad-gitlab.univ-grenoble-alpes.fr/OSUG/RESIF. The International Federation of Digital Seismograph Networks (FDSN) is available at https://www.fdsn.org. The Observatories and Research Facilities for European Seismology (ORFEUS) is available at https://www.orfeus-eu.org. The European Integrated Data Archive (EIDA) is available at https://www.orfeus-eu.org/data/eida/. The Résif seismic portal is available at http://seismology.resif.fr. A SeedLink server is accessible at rtserve.resif.fr. A complete catalog of Résif webservices is available at https://ws.resif.fr. Metadata from RAP network are available at https://ws.resif.fr/fdsnws/station/1/query?network=RA&level=network. The "RA" network DataCite XML is available at https://data.datacite.org/application/vnd.datacite.datacite+xml/10.15778/RESIF.RA. The DataCite search is available at https://search.datacite.org/works/10.15778/RESIF.RA. The Résif installation of YASMINE is available at https://yasmine.resif.fr. All websites were last accessed in March 2021.

Conflicts of Interest

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The authors first of all thank all the network operators and scientists who entrust their datasets to Résif-SI, and who work with us to ensure the highest possible level of data completeness, and data and metadata quality. The authors thank for their active participation in Résif-DC: Our colleagues from the Direction générale des services informatiques of the Université Grenoble Alpes (DGDSI, https://services-numeriques-personnels.univ-grenoble-alpes.fr) and the Unité Mixte de Service Grenoble Alpes Recherche - Infrastructure de Calcul Intensif et de Données (GRICAD; https://gricad.univ-grenoble-alpes.fr): G. Bruno, G. Enderlé, V. Louvet, and A. Monot; the initiator of the SUMMER project: J. Le Tanou, and the technical committee of the SUMMER storage infrastructure at Grenoble Alpes University (https://summer.univ-grenoble-alpes.fr); our colleagues from the Centre de données OSUG (OSUG-DC, https://www.osug.fr/missions/observation): B. Boutherin and R. Cailletau; our colleagues from the Centre de Calcul de l'IN2P3 (CC-IN2P3, https://cc.in2p3.fr/nos-services/traitement-de-donnees), and in particular J.-Y. Nief. Résif-SI relies on a national and international collaborative framework and feedback from multiple sources. The authors thank, in particular: the European Technical Committee (ETC) of European Integrated Data Archive (EIDA, https://www.orfeus-eu.org/data/eida/structure); Observatories and Research Facilities for European Seismology (ORFEUS-DC, https://www.orfeus-eu.org) for its support in the early days of Résif-DC, and in particular T. van Eck and R. Sleeman; the staff of Incorporated Research Institutions for Seismology (IRIS-DS, https://www.iris.edu/hq): T. Ahern, J. Carter, R. Casey, C. Trabant; the Résif-SI review committee (June 2014), for their precise and fruitful analysis: F. Genova (president), A. Ferreira, G. Guyot, J.-Y. Nief, C. Rondenay, C. Surace. The French seismological and geodetic network Résif-EPOS is a national facility cited in the national Roadmap for Research Infrastructures published by the Ministry of Higher Education, Research and Innovation. Résif-EPOS, which is coordinated by CNRS, receives funding, personnel and other support from all the Consortium members. Résif-SI is also supported by the System of Observation and Experimentation for Research and the Environment (SOERE) and by the Ministry of Ecological Transition. Résif-SI has benefited from the Résif-CORE project (11-EQPX-0040), funded by the national French funding program Investissements d'Avenir and managed by the French National Research Agency (ANR). The authors thank the two anonymous reviewers whose suggestions helped improve and clarify this article. This article is dedicated to the memory of Jocelyn Guilbert, former director of the "Laboratoire de Détection et de Géophysique" (Laboratory of Detection and Geophysics) of Commissariat à l'énergie atomique et aux énergies alternatives - Alternative Energies and Atomic Energy Commission (CEA), deceased on 21 August 2016. Jocelyn devoted a huge effort for the participation of CEA to Résif-SI and to the construction of Résif, in general. All websites were last accessed in March 2021.

References

Ahern, T., R. Casey, D. Barnes, R. Benson, T. Knight, and C. Trabant (2012). *Seed Reference Manual*, Incorporated Research Institutions for Seismology (IRIS), Washington, D.C., 220 pp., available at https://www.fdsn.org/publications (last accessed March 2021).

AlpArray Seismic Network (2015). AlpArray seismic network (AASN) temporary component, doi: 10.12686/ALPARRAY/Z3_2015.

Barruol, G., K. Sigloch, , and RHUM-RUM Group, and RESIF (2017). RHUM-RUM experiment, 2011-2015, code YV (Réunion Hotspot and Upper Mantle – Réunion's Unterer Mantel) funded by ANR, DFG, CNRS-INSU, IPEV, TAAF, instrumented by DEPAS, INSU-OBS, AWI and the Universities of Muenster, Bonn, La Réunion, doi: 10.15778/RESIF.YV2011.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010). ObsPy: A Python toolbox for seismology, *Seismol. Res. Lett.* **81,** no. 3, 530–533, doi: 10.1785/gssrl.81.3.530.

Bollard, P. (2021). Resif-delivery-stats-plotter, available at https://gitlab.com/resif/resif-delivery-stats-plotter (last accessed February 2021).

Brenguier, F., P. Kowalski, N. Ackerley, N. Nakata, P. Boué, M. Campillo, E. Larose, S. Rambaud, C. Pequegnat, T. Lecocq, *et al.* (2015). Toward 4D noise-based seismic probing of volcanoes: Perspectives from a large-N experiment on Piton de la Fournaise volcano, *Seismol. Res. Lett.* **87,** no. 1, 15–25, doi: 10.1785/0220150173.

Brenguier, F., D. Rivet, A. Obermann, N. Nakata, P. Boué, T. Lecocq, M. Campillo, and N. Shapiro (2016). 4-D noise-based seismology at volcanoes: Ongoing efforts and perspectives, *J. Volcanol. Geoth. Res.* **321,** 182–195, doi: 10.1016/j.jvolgeores.2016.04.036.

Brzev, S., C. Scawthorn, A. W. Charleson, L. Allen, M. Greene, K. Jaiswal, and V. Silva (2013). *GEM building taxonomy (Version 2.0)*, *Tech. Rep.*, GEM Foundation.

Casey, R. (1999). Networked data centers (NetDC), *Data Services Newsletter* 1**,** no. 1, available at https://ds.iris.edu/ds/newsletter/vol1/no1/225/networked-data-centers-netdc/ (last accessed September 2020).

Casey, R. (2016). Portable data collection center (PDCC) v3.8.1 user manual, Incorporated Research Institutions for Seismology (IRIS), available at https://ds.iris.edu/ds/nodes/dmc/software/downloads/pdcc (last accessed March 2021).

Clinton, J., W. Crawford, C. Evangelidis, P. Kaestli, P. Gueguen, H. Pedersen, J. Quinteros, J.-M. Saurel, R. Sleeman, and A. Strollo (2018). SERA deliverable 4.2: Report on metadata challenges and proposed solutions, *Tech. Rep.*, SERA, available at http://www.sera-eu.org/en/Dissemination/deliverables/ (last accessed March 2021).

Crawford, W. (2019). *OBS data/metadata proposal*, *Tech. Rep.*, FDSN, available at http://www.ipgp.fr/~crawford/documents.html (last accessed March 2021).

Crawford, W., R. Bouazzouz, and L. Arean (2019). Obsinfo, available at https://gitlab.com/resif/obsinfo (last accessed February 2021).

Creative Commons (2013). Attribution 4.0 International – CC BY 4.0, available at https://creativecommons.org/licenses/by/4.0 (last accessed October 2020).

DataCite (2019). DataCite metadata schema, available at https://schema.datacite.org (last accessed October 2020).

Dougherty, S. L., E. S. Cochran, and R. M. Harrington (2019). The Large-N seismic survey in Oklahoma (LASSO) experiment, *Seismol. Res. Lett.* **90,** no. 5, 2051–2057, doi: 10.1785/0220190094.

Evans, P., A. Strollo, A. Clark, T. Ahern, R. Newman, J. Clinton, H. Pedersen, and C. Pequegnat (2015). Why seismic networks need digital object identifiers, *Eos* **96,** doi: 10.1029/2015eo036971.

Flajolet, P., É. Fusy, O. Gandouet, and F. Meunier (2007). Hyperloglog: The analysis of a near-optimal cardinality estimation algorithm, *AofA'07: Proc. of The 2007 International Conf. on Analysis of Algorithms*, Nice, France, 17–22 June.

Folk, M., G. Heber, Q. Koziol, E. Pourmal, and D. Robinson (2011). An overview of the HDF5 technology suite and its applications, *Proc. of the EDBT/ICDT 2011 Workshop on Array Databases*, 36–47.

Geber, O. (2019). MORUMOTTO, available at https://github.com/IPGP/morumotto (last accessed October 2020).

GFZ and gempa GmbH (2008). The SeisComP seismological software package, doi: 10.5880/GFZ.2.4.2020.003.

GFZ and gempa GmbH (2008). SeedLink, available at https://www.seiscomp.de/doc/apps/seedlink.html (last accessed October 2020).

GFZ and gempa GmbH (2008). Waveform archives, available at https://www.seiscomp.de/doc/base/concepts/waveformarchives.html (last accessed October 2020).

Gimbert, F., U. Nanni, P. Roux, A. Helmstetter, S. Garambois, A. Lecointre, A. Walpersdorf, B. Jourdain, M. Langlais, O. Larman, *et al.* (2020). The RESOLVE project: A multi-physics experiment with a temporary dense seismic array on the Argentière Glacier, French Alps, available at https://arxiv.org/abs/2009.06321 (last accessed March 2021).

Goubier, O., and W. Crawford (2021). visualQC, available at https://gitlab.com/resif/visualqc (last accessed February 2021).

Grünthal, G. (ed.) (1998). *European Macroseismic Scale 1998 (EMS-98). Cahiers du Centre Européen de Géodynamique et de Séismologie 15*, *Tech. Rep.*, Centre Européen de Géodynamique et de Séismologie, Luxembourg, 99 pp., available at http://media.gfz-potsdam.de/gfz/sec26/resources/documents/PDF/EMS-98_Original_englisch.pdf (last accessed March 2021).

Gueguen, P. (2012). Experimental analysis of the seismic response of one base-isolation building according to different levels of shaking: Example of the Martinique earthquake (2007/11/29) Mw 7.3, *Bull. Earthq. Eng.* **10,** no. 4, 1285–1298, doi: 10.1007/s10518-012-9355-x.

Guéguen, P., F. Guattari, C. Aubert, and T. Laudat (2020). Comparing direct observation of torsion with array-derived rotation in civil engineering structures, *Sensors* **21,** no. 1, 142, doi: 10.3390/s21010142.

Hable, S., K. Sigloch, G. Barruol, S. C. Stähler, and C. Hadziioannou (2018). Clock errors in land and ocean bottom seismograms: High-accuracy estimates from multiple-component noise cross-correlations, *Geophys. J. Int.* **214,** no. 3, 2014–2034, doi: 10.1093/gji/ggy236.

Heinloo, A. (2018). fdsnws_scripts, available at https://geofon.gfz-potsdam.de/software/fdsnws_scripts (last accessed October 2020).

Hess, D., N. Falco, rsdeazevedo, damhuonglan, and K. Jacobs (2018). PIC-IRIS/PH5: V4.1.2, doi: 10.5281/ZENODO.841332.

Hutko, A. (2013). Web service fetch scripts, available at https://seiscode.iris.washington.edu/projects/ws-fetch-scripts (last accessed October 2020).

IN2P3/CNRS (1976). Centre de Calcul de l'IN2P3/CNRS – USR6402, available at https://cc.in2p3.fr (last accessed October 2020).

Institut de physique du globe de Paris (IPGP) (2007). 4G (2007-2025): EMSO-MoMAR, available at https://www.fdsn.org/networks/detail/4G_2007/ (last accessed September 2020).

International Federation of Digital Seismograph Networks (FDSN) (2013). FDSN web services, available at https://www.fdsn.org/webservices (last accessed October 2020).

International Federation of Digital Seismograph Networks (FDSN) (2019). StationXML schema 1.1, available at http://www.fdsn.org/xml/station/ (last accessed October 2020).

International Federation of Digital Seismograph Networks (FDSN) (2020). Network codes, available at https://www.fdsn.org/networks (last accessed October 2020).

IPGP (2008a). GNSS, seismic broadband and strong motion permanent networks in West Indies, doi: 10.18715/ANTILLES.WI.

IPGP (2008b). Seismic, deformation, gas, magnetic and weather permanent networks on Mount Pelée volcano and Martinique Island, doi: 10.18715/MARTINIQUE.MQ.

IPGP (2008c). Seismic, deformation, gas, magnetic and weather permanent networks on Piton de la Fournaise volcano and La Reunion Island, doi: 10.18715/REUNION.PF.

IPGP (2008d). Seismic, deformation, gas, magnetic and weather permanent networks on Soufrière volcano and Guadeloupe Island, doi: 10.18715/GUADELOUPE.GL.

Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre (EOST) (1982). GEOSCOPE, French global network of broad band seismic stations, doi: 10.18715/GEOSCOPE.G.

Incorporated Research Institutions for Seismology (IRIS) (2016). WebServiceShell, available at https://github.com/iris-edu/webserviceshell (last accessed October 2020).

ISO/IEC 19464 (2014). Information technology – Advanced message queuing protocol (AMQP) v1.0 specification, International Organization for Standardization, Geneva, Switzerland, available at https://www.amqp.org (last accessed March 2021).

Krischer, L., T. Megies, R. Barsch, M. Beyreuther, T. Lecocq, C. Caudron, and J. Wassermann (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discov.* **8,** no. 1, 014003, doi: 10.1088/1749-4699/8/1/014003.

Laigle, M., J. Lebrun, and A. Hirn (2007). Sismantilles 2 cruise, RV L'Atalante, doi: 10.17600/7010020.

Lorenzo, G. W. F., M. P. S. d'Avila, A. Deschamps, E. Bertrand, E. D. Mercerat, L. Foundotos, and F. Courboulex (2018). Numerical and empirical simulation of linear elastic seismic response of a building: The case of nice prefecture, *Earthq. Spectra* **34,** no. 1, 169–196, doi: 10.1193/042216eqs064m.

Loviknes, K., Z. Jeddi, L. Ottemöller, and T. Barreyre (2020). When clocks are not working: OBS time correction, *Seismol. Res. Lett.* **91,** no. 4, 2247–2258, doi: 10.1785/0220190342.

MacCarthy, J., O. Marcillo, and C. Trabant (2020). Seismology in the cloud: A new streaming workflow, *Seismol. Res. Lett.* **91,** no. 3, 1804–1812, doi: 10.1785/0220190357.

Massinon, B., and J. Plantet (1976). A large-aperture seismic network in France: Description and some results concerning epicenter location and upper-mantle anomalies, *Phys. Earth Planet. Int.* **12,** nos. 2/3, 118–127, doi: 10.1016/0031-9201(76)90041-8.

Masson, F., S. Auclair, D. Bertil, M. Grunberg, B. Hernandez, S. Lambotte, G. Mazet-Roux, L. Provost, J.-M. Saurel, A. Schlupp, *et al.* (2021). The transversal seismicity action RESIF: A tool to improve the distribution of the French seismicity products, *Seismol. Res. Lett.* doi: 10.1785/0220200353.

McNamara, D. E. (2004). Ambient noise levels in the Continental United States, *Bull. Seismol. Soc. Am.* **94,** no. 4, 1517–1527, doi: 10.1785/012003001.

McNamara, D. E., and R. I. Boaz (2010). PQLX: A seismic data quality control system description, applications, and users manual, *U.S. Geol. Surv. Open-File Rept. 1292*, 41 pp.

Megies, T., M. Beyreuther, R. Barsch, L. Krischer, and J. Wassermann (2011). ObsPy – What can it do for data centers and observatories? *Ann. Geophys.* **54,** no. 1, doi: 10.4401/ag-4838.

Michel, C., and P. Guéguen (2018). Interpretation of the velocity measured in buildings by seismic interferometry based on Timoshenko beam theory under weak and moderate motion, *Soil Dynam. Earthq. Eng.* **104,** 131–142, doi: 10.1016/j.soildyn.2017.09.031.

Michel, C., P. Guéguen, S. E. Arem, J. Mazars, and P. Kotronis (2009). Full-scale dynamic response of an RC building under weak seismic motions using earthquake recordings, ambient vibrations and modelling, *Earthq. Eng. Struct. Dynam.* 419–441, doi: 10.1002/eqe.948.

Michener, W. K. (2015). Ten simple rules for creating a good data management plan, *PLoS Comput. Biol.* **11,** no. 10, e1004525, doi: 10.1371/journal.pcbi.1004525.

MongoDB, Inc. (2009). MongoDB, available at https://www.mongodb.com (last accessed October 2020).

Observatories and Research Facilities for European Seismology (ORFEUS) (2020a). EIDA federator, available at https://www.orfeus-eu.org/data/eida/nodes/FEDERATOR (last accessed October 2020).

ORFEUS (2020b). EIDA webdc3, available at http://orfeus-eu.org/webdc3 (last accessed October 2020).

Pequegnat, C., P. Guéguen, D. Hatzfeld, and M. Langlais (2008). The French accelerometric network (RAP) and National Data Centre (RAP-NDC), *Seismol. Res. Lett.* **79,** no. 1, 79–89, doi: 10.1785/gssrl.79.1.79.

PostgreSQL Global Development Group (1996). PostgreSQL, available at https://www.postgresql.org (last accessed October 2020).

Quinteros, J., J. A. Carter, J. Schaeffer, C. Trabant, and H. A. Pedersen (2021). Exploring approaches for large data in seismology: User and data repository perspectives, *Seismol. Res. Lett.* doi: 10.1785/0220200390.

re3data.org (2016). RESIF seismic data portal, doi: 10.17616/R37Q06.

Résif (1995). RESIF-RAP French accelerometric network, doi: 10.15778/RESIF.RA.

Résif (2017). RESIF-RENAG French national geodetic network, doi: 10.15778/RESIF.RG.

Résif (2018). CEA/DASE broad-band permanent network in metropolitan France, doi: 10.15778/RESIF.RD.

Résif (2019a). Résif web services, available at http://ws.resif.fr (last accessed October 2020).

Résif (2019b). Webservice ph5-availability, available at http://ws.resif.fr/resifws/ph5-availability/1/ (last accessed October 2020).

Résif (2019c). Webservice ph5-dataselect, available at http://ws.resif.fr/resifws/ph5-dataselect/1/ (last accessed October 2020).

Résif-SI (2020). Networks citation page, available at http://seismology.resif.fr/#CMSConsultPlace:DOI_INVENTORY (last accessed October 2020).

Ronacher, A. (2010). Flask, available at https://flask.palletsprojects.com (last accessed October 2020).

Roult, G., J.-P. Montagner, E. Stutzmann, S. Barbier, and G. Guiveneux (1999). The GEOSCOPE program: Its data center, *Phys. Earth Planet. In.* **113,** nos. 1/4, 25–43, doi: 10.1016/s0031-9201(99)00024-2.

Saurel, J.-M., J. Schaeffer, and C. Péquegnat (2019). GATITO: Generic atomic library of tiny Objects: Library of metadata objects for solid earth science metadata, mostly enumerations, available at https://gitlab.com/resif/gatito (last accessed October 2020).

Shapiro, N., V. Douet, C. Pardo, C. Péquegnat, S. Barbier, C. Maron, M. Grunberg, and M. Schaming (2008). FOSFORE: Portal for distributions of French seismological data, *AGU Fall Meeting Abstracts*, Vol. 2008, S43D–1909.

Strollo, A., D. Cambaz, J. Clinton, P. Danecek, C. P. Evangelidis, A. Marmureanu, L. Ottemöller, H. Pedersen, R. Sleeman, K. Stammler, *et al.* (2021). EIDA: The European integrated data archive and service infrastructure within ORFEUS, *Seismol. Res. Lett* doi: 10.1785/0220200413.

Templeton, M. E. (2017). IRIS library of nominal responses for seismic instruments, doi: 10.17611/S7159Q.

Trabant, C. (2011). Ringserver – Generic packet ring buffer with network interfaces, available at https://github.com/iris-edu/ringserver (last accessed October 2020).

Trani, L., M. Koymans, M. Atkinson, R. Sleeman, and R. Filgueira (2017). WFCatalog: A catalogue for seismological waveform data, *Comput. Geosci.* **106,** 101–108, doi: 10.1016/j.cageo.2017.06.008.

Tridgell, A., and P. Mackerras (1996). The rsync algorithm, *Tech. Rep. TR-CS-96-05*, Australian National University, Department of Computer Science, available at http://rsync.samba.org (last accessed March 2021).

Université Grenoble Alpes (2016). SUMMER: Stockage Unifié Mutualisé Massif Evolutif et Réparti, available at https://summer.univ-grenoble-alpes.fr (last accessed October 2020).

Vallée, M., J. Charléty, A. M. G. Ferreira, B. Delouis, and J. Vergoz (2011). SCARDEC: A new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body-wave deconvolution, *Geophys. J. Int.* **184,** no. 1, 338–358, doi: 10.1111/j.1365-246x.2010.04836.x.

Voigt, C., C. Förste, H. Wziontek, D. Crossley, B. Meurers, V. Pálinkáš, J. Hinderer, J.-P. Boy, J.-P. Barriot, and H. Sun (2016). Report on the data base of the international geodynamics and earth tide service (IGETS), *Tech. Rep.* doi: 10.2312/GFZ.B103-16087.

Volcke, P., R. Bouazzouz, J. Schaeffer, and C. Péquegnat (2013). ResifDataTransfer, available at https://gitlab.com/resif/resif-data-transfer (last accessed October 2020).

Volcke, P., J. Schaeffer, and C. Péquegnat (2012). seedtree5, available at https://gitlab.com/resif/seedtree5 (last accessed October 2020).

Weertman, B. (2010). Web services at the DMC, *Data Services Newsletter* 12**,** no. 3, available at http://ds.iris.edu/ds/newsletter/vol12/no3/44/web-services-at-the-dmc (last accessed October 2020).

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, *et al.* (2016). The FAIR guiding principles for scientific data management and stewardship, *Sci. Data* **3,** no. 1, doi: 10.1038/sdata.2016.18.

Wolyniec, D., J. Schaeffer, J.-M. Saurel, and C. Péquegnat (2019). AROL: Atomic response objects library containing metadata descriptions of earth science observation instruments, available at https://gitlab.com/resif/arol (last accessed October 2020).

Zaccarelli, R., D. Bindi, A. Strollo, J. Quinteros, and F. Cotton (2019). Stream2segment: An open-source tool for downloading, processing, and visualizing massive event-based seismic waveform datasets, *Seismol. Res. Lett.* **90,** no. 5, 2028–2038, doi: 10.1785/0220180314.

Appendix

Data access details and examples

In this appendix, we present the different ways for end users to access Résif-SI data and metadata.

Real time: SeedLink. Résif-DC real-time data service (SeedLink server, GeoForschungsZentrums [GFZ], and gempa GmbH, 2008) provides access to data for the following permanent networks: FR, RD, RA, MT, G, WI, ND, CL, MQ, PF, and GL. Data from temporary campaigns are usually not distributed in real-time, unless acquired in the context of a postseismic campaign for which Résif does not accept any data embargo time. The lifetime of the near-real-time data in the internal memory buffer is about 3 hr. By convention, Résif-DC does not distribute any raw data for which the corresponding metadata are not available through the International Federation of Digital Seismograph Networks (FDSN) webservice station.

The SeedLink service is available at: rtserve.resif.fr (TCP port 18000).

Webservices. Webservices are the base delivery mechanism for most Résif data and metadata. Résif-DC exposes the four webservices standardized by the FDSN (FDSN, 2013), an European Integrated Data Archive (EIDA) webservice common to all EIDA DCs, as well as other more specific webservices (Résif, 2019a).

FDSN webservices. The FDSN webservices are characterized by common and standardized interfaces, and, thus, constitute a homogeneous means of access to data, regardless of the data center involved. Those services are all accessible with the prefix https://ws.resif.fr/fdsnws/ (Table A1).

Examples:

- All Résif stations, XML formatted: https://ws.resif.fr/fdsnws/station/1/query?level=station (last accessed March 2021).
- One hour of data for network ZH (516 MB): https://ws.resif.fr/fdsnws/dataselect/1/query?network=ZH&starttime=2003-06-01T00:00:00&endtime=2003-06-01T01:00:00 (last accessed March 2021).
- Available data from network G, station SSB2, channel BHZ for a given time interval: https://ws.resif.fr/fdsnws/availability/1/query?network=G&station=SSB&channel=BHZ&start=2010-02-23T00:00:00&end=2010-03-23T00:00:00 (last accessed March 2021).

FDSN webservices can be accessed by standard tools and libraries, such as ObsPy (Beyreuther *et al.*, 2010; Megies *et al.*,

TABLE A1
FDSN Webservices Provided by Résif-DC, with Base URL https://ws.resif.fr/fdsnws/ (last accessed March 2021)

Webservice/Version	Comment	Format
station/1	Metadata inventory	StationXML or text
dataselect/1	Time-series data	miniSEED
availability/1	Data inventory	Text, JSON, geocsv
event/1	Event parameters	quakeML

Note that the event webservice is operated by Réseau National de Surveillance Sismique (RéNaSS) – National Seismic Monitoring Network.

TABLE A2
EIDAWS Webservices Provided by Résif-DC, with Base URL https://ws.resif.fr/eidaws/ (last accessed March 2021)

Webservice/Version	Comment	Format
wfcatalog/1	Data quality and metrics	JSON

2011; Krischer *et al.*, 2015). As an example, we provide in the Appendix an example Python script to retrieve data and metadata through the FDSN webservices.

EIDA webservices. Résif-DC exposes the EIDAWS wfcatalog webservice (Trani *et al.*, 2017). This webservice provides detailed metrics on waveform data (e.g., gaps, overlaps, root mean square [rms], Standard for Exchange of Earthquake Data [SEED] data quality flag). The WFCatalog interface can be used for data discovery, as it supports range filtering on all available metrics. The EIDAWS webservice is accessible with the prefix https://ws.resif.fr/eidaws/ (Table A2).

Examples:

- Waveform metadata documents with a daily granularity from network CL, between 1 January 2018 T00:00:00 and 10 January 2018 T00:00:00, in JSON format: https://ws.resif.fr/eidaws/wfcatalog/1/query?network=CL&include=sample&start=2018-01-01&end=2018-01-10 (last accessed March 2021).

Other webservices. Résif-DC has implemented additional webservices, whose functionalities have been inspired by the services provided by the Incorporated Research Institutions for Seismology (IRIS) data services (Weertman, 2010) or needed by the French users of Résif-SI. The former are accessible with the prefix https://ws.resif.fr/resifws/, while the latter are accessible with the prefix https://ws.resif.fr/resifsi/. Tables A3 and A4 show complete lists of these two groups of services.

Examples:

- five minutes of data for station PYTO, channel HN2, with demean and response deconvolution, ASCII format: https://ws.resif.fr/resifws/timeseries/1/query?net=RA&station=PYTO&cha=HN2&loc=02&demean&correct&start=2017-11-02T13:35:00&end=2017-11-02T13:40:00&format=ascii (last accessed March 2021);
- interactive plot of five minutes of data for station PYTO, channel HN2, location code 02, with demean and response deconvolution: https://ws.resif.fr/resifws/timeseriesplot/1/query?net=RA&station=PYTO&cha=HN2&loc=02&demean&correct&start=2017-11-02T13:35:00&end=2017-11-02T13:40:00&iplot (last accessed March 2021);
- plot of instrument response for station CIEL, network FR, channel HHE, location code 00: https://ws.resif.fr/resifws/evalresp/1/query?net=FR&sta=CIEL&loc=00&cha=HHE&

TABLE A3
RESIFWS Webservices Provided by Résif-DC, with Base URL https://ws.resif.fr/resifws/ (last accessed March 2021)

Webservice/Version	Comment	Format
timeseries/1	Time-series data	miniSEED, sac, text, png
timeseriesplot/1	Time-series data	jpeg or png
ph5-dataselect/1	Time-series data	miniSEED, sac, geocsv (converted from PH5)
ph5-availability/1	Data inventory	text, json, geocsv
evalresp/1	Response information evaluated from Résif metadata	text or png
resp/1	Channel response information	resp (text)
sacpz/1	Channel response information	sacpz (text)
statistics/1	Data download statistics from Résif-DC	text

Note that ph5-dataselect and ph5-availability comply to the FDSN webservices specification (FDSN, 2013) and might, therefore, be migrated in the future to a standard URL. For up-to-date information on Résif webservice URLs, see Data and Resources.

TABLE A4
Webservices Provided by Résif-DC for Data Providers, with Base URL https://ws.resif.fr/resifsi/ (last accessed March 2021)

Webservice/Version	Comment	Format
transaction/1	Status of an data or metadata integration transaction; this service is for Résif data providers	text, json
orphanfile/1	Data (SDS files) without metadata in Résif archive; this service is for Résif data providers	text
assembleddata/1	Historical and validated collections of the permanent accelerometric network (1995–2013)	tar.gz

time=2017-01-01T00:00:00&format=plot (last accessed March 2021);

- response (RESP format) for station RUSF, north channel, location code 06: https://ws.resif.fr/resifws/resp/1/query?net=FR&sta=RUSF&loc=06&cha=??N&time=2017-01-01T00:00:00&nodata=404 (last accessed March 2021);
- data availability for an experiment stored as PH5 archive: http://ws.resif.fr/resifws/ph5-availability/1/extent?net=3C (last accessed March 2021);
- miniSEED data from a network stored as PH5 archive: http://ws.resif.fr/resifws/ph5-dataselect/1/query?net=3C&starttime=2019-12-09T00:00:00&endtime=2019-12-09T00:00:10 (last accessed March 2021);
- quantity of data stored for year 2020 by network FR: http://ws.resif.fr/resifws/statistics/1/query?request=storage&net=FR&type=validated&year=2020 (last accessed March 2021);
- quantity of data shipped via dataselect for network WI: http://ws.resif.fr/resifws/statistics/1/query?request=send&media=dataselect&net=WI (last accessed March 2021).

Specificity of OBS data and metadata. Ocean-bottom seismometer (OBS) data are archived both in time-corrected format (quality label "Q") and in original, unmodified format (quality label "D"). The following request (with no quality parameter specified) delivers both corrected (Q label) and raw data (D label):

```
wget –http-user=USER –http-password=PASSWORD "https://ws.resif.fr/fdsnws/dataselect/1/queryauth?network=Z3&station=A4*&starttime=2017-08-01&endtime=2017-08-01T01:00:00" -0 mydata.miniseed
```

By specifying the quality=Q parameter, only corrected data (Q label) is retrieved:

```
wget –http-user=USER –http-password=PASSWORD "https://ws.resif.fr/fdsnws/dataselect/1/queryauth?network=Z3&station=A4*&starttime=2017-08-01&endtime=2017-08-01T01:00:00&quality=Q" -O mydata.miniseed
```

With quality = D, only raw data (uncorrected) is delivered.

The StationXML metadata associated with the OBS are progressively refined with information on time corrections, instrument positioning uncertainties, water column depth, and precise instrument description (sensor, data logger). A typical example of OBS metadata at Résif-DC is provided in the code block above.


```
<Station code="A401A" startDate="2017-06-22T01:59:00" endDate="2018-02-16T00:01:00" restrictedStatus="closed" resif:alternateNetworkCodes="Z32015">
  <Comment><Value> {"clock_correction": {
    "linear_drift": {
      "time_base": "Seascan MCXO, 1e-8 nominal drift",
      "reference": "GPS", "start_sync_instrument": 0,
      "start_sync_reference": "2017-06-21T06:36:00Z",
      "end_sync_reference": "2018-02-16T01:44:00.146Z",
      "end_sync_instrument": "2018-02-16T01:44:00.00Z"
    }
  }}</Value></Comment>
  <Latitude minusError="4.5e-05" plusError="4.5e-05" unit="DEGREES">42.647</Latitude>
  <Longitude minusError="6.12e-05" plusError="6.12e-05" unit="DEGREES">5.0198</Longitude>
  <Elevation minusError="10" plusError="10" unit="METERS">-1619.0</Elevation>
  <Site><Name>N/W Mediterranean Sea</Name></Site>
  <Vault><Name>Sea floor</Name></Vault>
  <WaterLevel minusError="10" plusError="10" unit="METERS">1619.</WaterLevel>
</Station>
```


Authors and Affiliations

Catherine Péquegnat: University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France, https://orcid.org/0000-0002-9956-9971;

Jonathan Schaeffer: University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0002-2447-7394; *Claudio Satriano: Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0002-3039-2530; **Helle Pedersen:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0003-2936-8047; **Jérôme Touvier:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France; **Jean-Marie Saurel:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0003-1458-4193; **Marie Calvet:** Institut de Recherche en Astrophysique et Planétologie, Université Toulouse III Paul Sabatier, CNRS, CNES, Toulouse, France, https://orcid.org/0000-0001-9660-5079; **Laurent Stehly:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0002-1854-7157; **Patrick Arnoul:** CEA/DAM/DIF, Arpajon, France; **Philippe Bollard:** University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0001-9156-7819; **Olivier Dewée:** Division Technique de l'INSU, Meudon, France; **Fabien Engels:** Université de Strasbourg, CNRS, EOST UMS, Strasbourg, France; **Marc Grunberg:** Université de Strasbourg, CNRS, EOST UMS, Strasbourg, France, https://orcid.org/0000-0002-1307-7790; **Constanza Pardo:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0003-1601-8718; **David Wolyniec:** University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France; **Grégory Arneodo:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France; **Rima Bouazzouz:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France; **Olivier Geber:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France; **Coralie Aubert:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France, https://orcid.org/0000-0003-3198-7402; **Mickaël Langlais:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France, https://orcid.org/0000-0003-0597-8289; **Sandrine Roussel:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France; **Florent Brenguier:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0001-8684-7613; **Wayne C. Crawford:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0002-3260-1826; **Arnaud Lemarchand:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0002-4234-3787; **Emeline Maufroy:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0002-0995-5597; **Serge Olivier:** CEA/DAM/DIF, Arpajon, France; **Diane Rivet:** Université Côte d'Azur, CNRS, IRD, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France, https://orcid.org/0000-0003-4697-4460; **Martin Vallée:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0001-8049-4634; **Jérôme Vergne:** Université de Strasbourg, CNRS, IPGS, Strasbourg, France, https://orcid.org/0000-0003-1731-9360; **Jérôme Chèze:** Université Côte d'Azur, CNRS, IRD, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France; **Éric Debayle:** Laboratoire de Géologie de Lyon: Terre, Planètes, Environnement, CNRS, École Normale Supérieure de Lyon, Université de Lyon, Université Claude Bernard Lyon 1, Villeurbanne, France, https://orcid.org/0000-0002-9438-7060; **Bertrand Delouis:** Université Côte d'Azur, CNRS, IRD, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France, https://orcid.org/0000-0003-1147-5872; **Philippe Guéguen:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France, https://orcid.org/0000-0001-6362-0694; **Sophie Lambotte:** Université de Strasbourg, CNRS, IPGS, Strasbourg, France, https://orcid.org/0000-0003-1078-3419; **Antoine Mocquet:** Laboratoire de Planétologie et de Géodynamique, Université de Nantes, Université d'Angers, Nantes, France, https://orcid.org/0000-0002-7832-3905; **Anne Paul:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France, https://orcid.org/0000-0002-5973-1786; **Fabrice Peix:** Université Côte d'Azur, CNRS, IRD, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France; **François Schindelé:** CEA/DAM/DIF, Arpajon, France; **Nikolai Shapiro:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France, https://orcid.org/0000-0002-0144-723X; **Éléonore Stutzmann:** Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France, https://orcid.org/0000-0002-4348-7475; **Pierre Volcke:** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France and **Andrea Walpersdorf** University of Grenoble Alpes, University of Savoie Mont Blanc, CNRS, IRD, University of Gustave Eiffel, ISTerre, Grenoble, France University of Grenoble Alpes, Irstea, CNRS, IRD, Grenoble, France, https://orcid.org/0000-0003-1777-9172

Manuscript received 21 October 2020
Published online 7 April 2021