

Réf.
Version
Date
DOI

IMN/IDM/ISI/FM/16-024
1.0
13th May 2016
<http://dx.doi.org/10.13155/44515>



Ifremer

**DOIs for ocean data
General principles and selected examples
(Argo, French cruises)**

Table of content

1. General principles for data DOI	3
1.1. DOI definition	3
1.2. Why to attribute DOIs to research data?	3
1.3. Data access services from a Landing Page	3
1.4. Should the entity that assign DOIs be responsible for data preservation?.....	3
1.5. How to cite data in a publication?	4
1.6. Where to cite a data DOI in a publication?	4
1.7. How to attribute data DOIs through DataCite?	4
1.8. What is the cost of an individual DOI?	5
1.9. What granularity to choose?.....	5
1.10. How to manage different versions of the same data set.....	5
1.11. How to manage dynamic data sets.....	5
1.12. What complementarity with databases?	6
1.13. Can a DOI be deleted?.....	6
1.14. Where to find the DOIs assigned by DataCite?	6
2. Examples of DOI attribution policies.....	7
2.1. Argo.....	7
2.1.1. Organization	7
2.1.2. Content of the global Argo DOI LandingPage (“tagless”).....	8
2.1.3. Content of a specific snapshot of Argo DOI LandingPage (“tagged”).....	10
2.2. The DOI of French oceanographic cruises.....	11
2.2.1. Objectives for the cruises’ DOIs	11
2.2.2. Organization	11
2.2.3. Series of cruises.....	11
2.2.4. Content of a sample Landing Page a cruise DOI.....	11

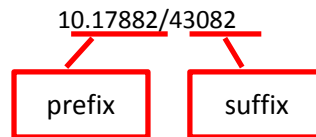
1. General principles for data DOI

1.1. DOI definition

A DOI (Digital Object Identifier) is an allocation of unique identifier. Generally used to identify scientific publications, a DOI can be attributed to any physical, numerical or abstract resource.

DOI are assigned by authorized entities. Some of these entities, such a DataCite are specialized in DOI for scientific datasets.

A DOI is a character string, including a prefix attributed to an entity and a suffix assigned by the entity:



An entity may have distinct prefixes. For example, Ifremer uses 5 different prefixes for 5 distinct projects. The suffix contains letter and/or figures. There is no limitation on the size of suffixes.

A DOI can be mentioned starting with a « DOI : » or preferably with the URL of the DOI solver « <http://dx.doi.org/> » or « <http://doi.org/> » (Example : <http://doi.org/10.17882/43082>). Such an URL becomes the official URL of the resource on the web.

A click on a DOI sends a request to the DOI entity (such as DataCite) the redirects transparently the user toward a web page called the « LandingPage » (LP). The LP describes the resource associated to the DOI, with a series of metadata and a link to data access.

1.2. Why to attribute DOIs to research data?

A DOI is attributed to a dataset to address several objectives:

Rationalize Data citation: if a resource editor changes, the URL of the landing page of the DOI will be updated. The bibliographic citation of the DOI remains valid.

Simplify data access: the landing page of a DOI offers a direct access to data through an HTTP/FTP link. As the data associated to a publication is publicly available, a reader can check the publication against the DOI data, increasing the publication credibility.

Provide traceability on data usage: the publications that cite a data set with a DOI are easy to track in bibliographic surveys. A comprehensive bibliographic survey will increase the notability of the scientist or infrastructure that published the data set with a DOI.

1.3. Data access services from a Landing Page

Ideally, data access is provided by a direct FTP or HTTP URL listed in the DOI's Landing Page. The data access may also be a simple email address to contact the data producer, or a contact web form.

Following European recommendations, free and direct access should be provided for data coming from public funding activities.

Some data providers propose an embargo on data access; a DOI may be assigned to data with restricted access, limited in time, during for instance the pair e=review period of a scientific publication.

1.4. Should the entity that assign DOIs be responsible for data preservation?

An entity may assign a DOI to datasets hosted in another entity. But this delegation is not safe, as the Landing page may link to obsolete or invalid resource if the relation between the two entity weakens years after years.

1.5. How to cite data in a publication?

DataCite suggests to cite a data set in a format similar to publication citations. Example of citation format:

Creator (PublicationYear): Title. Publisher. Identifier

Or

Creator (PublicationYear): Title. Version. Publisher. ResourceType. Identifier

1.6. Where to cite a data DOI in a publication?

Some journals such as “Science” ask the authors to cite the data in the Acknowledgments section. In some other journals, some authors cite their data in the “References” section of their article. Then, this section will display references of publications and datasets in undifferentiated manner.

Example of data citation in a publication

References

Berry J, Björkman O (1980) Photosynthetic response and adaptation to temperature in higher plants. *Annu Rev Plant Physiol* 31:491–543
Bindi M, Fibbi L, Gozzini B, Orlandini S, Miglietta F (1996) Modelling the impact of future climate scenarios on yield and yield variability of grapevine. *Clim Res* 7:213–224
Bindi M, Fibbi L, Miglietta F (2001) Free Air CO₂ Enrichment (FACE) of grapevine (*Vitis vinifera* L.): II. Growth and quality of grape and wine in response to elevated CO₂ concentrations. *Eur J Agron* 14:145–155

Kenny GJ, Harrison PA (1992) The effects of climate variability and change on grape suitability in Europe. *J Wine Res* 3:163–183
Kistler R et al (2001) The NCEP/NCAR 50-year reanalysis: monthly-means CD-ROM and documentation. *Bull Am Meteorol Soc* 82:247–267

Lautenschlager M et al (2009a) Climate Simulation with CLM, Climate of the 20th Century run no.1, Data Stream 3: European region MPI-M/ MaD. World Data Center for Climate. doi:10.1594/WDC/CLM_C20_1_D3
Lautenschlager M et al (2009b) Climate Simulation with CLM, Climate of the 20th Century run no.2, Data Stream 3: European region MPI-M/ MaD. World Data Center for climate. doi:10.1594/WDC/CLM_C20_2_D3

Lautenschlager M et al (2009c) Climate Simulation with CLM, Climate

Figure 1 : citation of a dataset in a publication

1.7. How to attribute data DOIs through DataCite?

There are several data publishing repositories¹ in Europe that propose free publication of research data with DOIs.

An entity that want to implement a DOI attribution mechanism for its data holdings through DataCite will first subscribe a contract with the national representative of DataCite (INIST/CNRS in France).

DataCite then attributes one or many prefixes to the institution for its future DOIs.

To activate a DOI, the entity will post the request to DataCite either through a web form or an API – an XML flux that contains a set of metadata with 5 mandatory ones:

- Identifier: the DOI
- Creator: name of the main scientists involved in the data production or a list collectivity of authors (ex: project, name of organisms)
- Title: the resource title
- Publisher: name of the editing, publishing, distributing entity (ex : World Data Center for Climate)
- PublicationYear: year of public delivery of the resource. This year is a significant date for the resource.

These metadata should be posted with the URL of the LandingPage. Whenever needed, the metadata and the URL can and should be updated.

¹ Seanoe (<http://www.seanoe.org>), Pangaea (<https://www.pangaea.de>), Zenodo (<https://zenodo.org>), ...

More metadata² may complement these 5 mandatory metadata, to provide a comprehensive description of the dataset.

1.8. What is the cost of an individual DOI?

The cost of attribution of a DOI differs between agencies. In France, a scientific institution should subscribe an annual contract of 180€ for an unlimited stock of DOIs.

1.9. What granularity to choose?

There is no definitive answer to that question; it is a choice that will depend on data and their usage by scientists. The DOIs main objective is to facilitate and improve the reliability of data citation in scientific publications: a strategy should be defined based on the level of aggregation of data that are generally used in a publication.

If a DOI is assigned at a measurement level and if the related publication manages thousands of these measurements, it will be difficult to cite properly the thousand associated DOIs. In that case, DOIs are probably not properly assigned.

It is likely that for scientific publications exploiting different levels of the same bunch of data, DOIs should be assigned at these specific distinct levels. In the Landing Page of such a DOI it is possible to provide links between the different DOIs so that the reader will navigate from one level to another. In the metadata posted to DataCite when assigning this DOI, the relationship with related DOI is specified using roles such as "isPartOf" or "hasPart".

1.10. How to manage different versions of the same data set

There are no strong rules on the management of versions of a dataset. The management may be balanced for large volume data sets by the cost of storage versus the interest to access a series of previous versions of these datasets.

If different versions of a dataset should continue to be accessible independently and specifically cited (e.g. with a list of different authors), a specific DOI should be assigned for each version.

It is possible to provide access to different versions of a distinct dataset in the LandingPage of a DOI. Seanoe data publisher provides in this case two choices:

1. All versions of the data set are directly available
2. Access to old versions is restricted: previous versions are accessible via a request to a service desk. It is then possible to ensure the reproducibility of an experiment based on an obsolete version, but access to an obsolete version of the data set is intentional and not accidental.

Data	File	Size	Format	Processing	Access
	Version 1	38 KB	CSV	Raw data	Restricted access
	Version 2	38 KB	CSV	Raw data	Open access

Figure 2 : Example of a series of versions of a dataset in a DOIs LandingPage. The latest version 2 is directly available; the obsolete version 1 is available on demand only.

1.11. How to manage dynamic data sets

Dynamic data sets receive regularly new data or data updates. The Research Data Alliance (RDA) published a recommendation³ to enable researchers to identify and cite them.

² <https://schema.datacite.org/>

³ https://rd-alliance.org/system/files/documents/RDA-DC-Recommendations_151020.pdf

1.12. What complementarity with databases?

Some data sets are continuously managed in databases. Is there an interest to export the database to publish and quote it with a specific DOI?

The allocation of DOI from dedicated warehouses is indeed an answer to the need of reproducibility and accessibility.

To get a DOI, the dataset extracted from a database is stored in a dedicated data warehouse. The dataset is fixed. When quoted in a publication, a user who question the validity of a result, can get the exact version of the data used in the publication. This need for reproducibility is generally not covered by existing databases; they do not systematically record changes in data.

The export and publication of a set of data from a database can also simplify access by offering a download in the form of one or more files available in one click from the LandingPage. In principle, a DOI may be assigned to a database. In this case, if a user discovers the existence of data via a publication, and if he wishes to use these own work, he will first have to deal with the interface of the database before getting access to the information.

Many databases do not retain the list of authors of the data they aggregate. They do not manage properly the authors citation on specific sub-datasets.

For example, here is an exported dataset from the Quadrigé database to Seanoe to be quoted in a publication:

Pouvreau Stephane, Maurer Daniele, Auby Isabelle, Lagarde Franck, Le Gall Patrik, Cochet Hélène, Bouquet Anne-Lise, Geay Amélie, Mille Dominique (2016). VELYGER Database: The Oyster Larvae Monitoring French Project. Seanoe. <http://doi.org/10.17882/41888>

Quadrigé is an environmental monitoring relational database. A direct query of the database will not credit individual scientists. A specific dataset exported from Quadrigé database, registered in Seanoe with a DOI will document and credit the individual scientists who contributed to this dataset.

In this example, the authors are personally cited, a credit for their contribution to the creation of the dataset. The data are available in open access with one click from the LandingPage to the data file.

A direct access from Quadrigé database would not cite the list of authors. If a user wants more data, the direct link to Quadrigé database web interface is provided in the Landing Page.

1.13. Can a DOI be deleted?

A DOI cannot be deleted. Contractually, an organization that assigns a DOI is committed to maintaining online LandingPage for each DOI. If a data set becomes unavailable, the LandingPage must then present to users the reasons for the access removal to the data.

Scientific publications may be cited decades after their publication. The allocation of DOI data should enable a reliable and durable quote on the same timescales.

1.14. Where to find the DOIs assigned by DataCite?

The metadata of resources with a DataCite DOI are searchable on

- <https://search.datacite.org/>

2. Examples of DOI attribution policies

2.1. Argo

2.1.1. Organization

Argo data are collected and disseminated by Argo GDACs (Global Data Assembly Centers) through their FTP sites. The available data from these FTP sites are continuously changing: data is added and updated every day.

To allow reproducibility of studies with Argo data, a snapshot of the entire data set is preserved every month. The snapshot contains all the Argo data available at the time of the snapshot creation. The one-month period between two snapshots was decided by the Scientific Committee of Argo; within a given month, changes to the whole dataset are not significant.

Initially, according to the suggestion of DataCite to include dynamic data, a DOI was set to describe the overall data set: he proposed access via FTP sites GDAC.

Argo (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). Ifremer. <http://dx.doi.org/10.12770/1282383d-9b35-4eaa-a9d6-4b0c24c0cfc9>

In addition, specific DOIs were assigned to each monthly snapshot:

Argo (2016). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) - Snapshot of Argo GDAC as of February, 8th 2016. Ifremer. <http://dx.doi.org/10.12770/9d8ac2dc-6f4d-4379-8df2-714cab4a9ae7>

Argo (2016). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) - Snapshot of Argo GDAC as of January, 8th 2016. Ifremer. <http://doi.org/10.12770/9fcd650a-5c41-47a7-8b56-23177962e81a>

...

In March 2016, to satisfy the Argo group that did not want to inflate the number of DOIs and to follow the new recommendations of the Research Data Alliance (RDA), a **new single Argo DOI** was published using Seanoe (a data publisher). This unique DOI quotes either the global data set or a specific snapshot. Each monthly snapshot is uploaded in Seanoe that assigns a URL and a key. The key 42350 for example, was assigned to the snapshot 2016-02-08.

The citation of the whole data set is performed by citing the new DOI without parameters:

Argo (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). Seanoe. <http://doi.org/10.17882/42182>

The DOI LandingPage then presents general information detailed in paragraph §1.7.

The citation of a specific snapshot is done by adding the key preceded by the # character to the DOI:

Argo (2016). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) - Snapshot of Argo GDAC of February, 8th 2016. Seanoe. <http://doi.org/10.17882/42182#42350>

This DOI Landing Page then provides information about the specific snapshot (Figure 5).

This new single Argo DOI offers an easier identification of publications as they simply cite the unique DOI. The snapshots are archived in Seanoe; this is a guarantee for long term.

The homogenization of URL data access allows the calculation of downloads statistics that are automatically provided to the authors once a year. This service to data providers is available on Seanoe.

The solution developed for publishing the Argo data in Seanoe can be implemented without additional development for any type of marine data.

2.1.2. Content of the global Argo DOI LandingPage (“tagless”)

General data (Figure 3-1)

In the main part of the page, the dataset is presented with a series of general metadata (title, author, publication date, description, Creative Commons license, ...).

Link User's manual (Figure 3-2)

The link to Argo User's manual with the DOI assigned to that document.

List of snapshots (Figure 3-3)

By default, the last three snapshots are listed. A link (Figure 3-4) displays all available snapshots. A link to download each snapshot (Figure 3-13).

Link to FTP sites GDAC (Figure 3-5)

The link to the ftp servers of the Argo GDAC (Argo global data assemble center).

Suggested citation (Figure 3-6)

The citation proposal is built according to a format suggested by DataCite:

“Creator (PublicationYear): Title. Publisher. Identifier”

Geographical area (Figure 3-7)

The geographic coverage of the dataset.

Export metadata (Figure 3-8)

The link to download metadata of the DOI in the RIS format. This format allows automatic import in bibliographic management tools (e.g. Endnote).

List of citing publications (Figure 3-9)

The list of publications documents that cite the data DOI. The documents may also have their individual DOI. A regular monitoring is performed by Simer on publications which cite datasets deposited in Seanoe DOI. For Argo, this monitoring is made by Megan Scanderbeg of the University of California San Diego. Seanoe queries the list of DOI and identified publications every night to build and update the list of publications.

List of associated data sets (Figure 3-10)

The list of related datasets that may also have their own DOI.

Link to social networks (Figure 3-11)

This link provides an automatic reporting of the dataset to social networks, through its DOI.



1

Argo float data and metadata from Global Data Assembly Centre (Argo GDAC)

Publication date 2000-09-12
 Author(s) Argo
 DOI [10.17882/42182](https://doi.org/10.17882/42182)
 Publisher SEANOE
 Keyword(s) float, Argo, global ocean observing system, ocean circulation, in-situ, ocean pressure, sea water salinity, sea water temperature, multi-year, weather climate and seasonal observation, global-ocean, Installations de suivi environnemental
 Abstract Argo is a global array of 3,000 free-drifting profiling floats that measures the temperature and salinity of the upper 2000 m of the ocean. This allows, for the first time, continuous monitoring of the temperature, salinity, and velocity of the upper ocean, with all data being relayed and made publicly available within hours after collection. The array provides 100,000 temperature/salinity profiles and velocity measurements per year distributed over the global oceans at an average of 3-degree spacing. Some floats provide additional bio-geo parameters such as oxygen or chlorophyll. All data collected by Argo floats are publicly available in near real-time via the Global Data Assembly Centers (GDACs) in Brest (France) and Monterey (California) after an automated quality control (QC), and in scientifically quality controlled form, delayed mode data, via the GDACs within six months of collection.
 Licence
 Utilisation A user of Argo data is expected to read and understand this manual and the documentation about the data contained in the "attributes" of the NetCDF data files, as these contain essential information about data quality and accuracy. A user should acknowledge use of Argo data in all publications and products where such data are used, preferably with the DOI and following standard sentence: "These data were collected and made freely available by the international Argo project and the national programs that cooperate with it."
 User's manual <http://dx.doi.org/10.13155/29825>

3

File	Size	Format	Processing	Access	Key
2016-03-08	9 GB	NC, NetCDF	Quality controlled data	Open access	42382
2016-02-08	8 GB	NC, NetCDF	Quality controlled data	Open access	42350
2016-01-08	8 GB	NC, NetCDF	Quality controlled data	Open access	42349

13

Data <ftp://ftp.ifremer.fr/ifremer/argo>
<ftp://usgodae.org/pub/outgoing/argo>

5

[More snapshots](#)

4

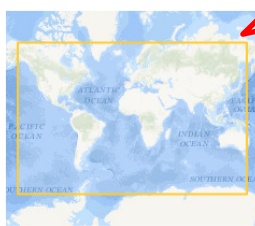
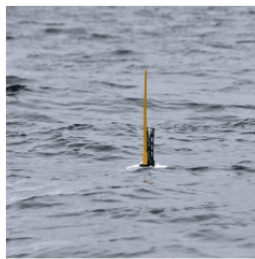
How to cite

Argo (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE. <http://doi.org/10.17882/42182>

6

[Top of the page](#) ↑

Click to download the data



7

Download metadata
 TXT, RIS, XLS

8

Related publications

[Fousiya T. S., Parekh Anant, Srinaseelan C. \(\). Interannual variability of upper ocean stratification in Bay of Bengal: observational and modeling aspects. Theoretical and Applied Climatology, -.](#)

[Kawakami Yuma, Sugimoto Shusaku, Suga Toshio \(2016\). Inter-annual zonal shift of the formation region of the lighter variety of the North Pacific Central Mode Water. Journal of Oceanography, 72\(2\), 225-234.](#)

9

Related datasets

[Feucher Charlene, Maze Guillaume, Mercier Herle \(2016\). 2000-2014 climatology of the North Atlantic permanent pycnocline properties. Seanoe.](#)

10

Share

11

Figure 3 : landing Page of the global Argo DOI

2.1.3. Content of a specific snapshot of Argo DOI LandingPage ("tagged")

When the DOI Argo is called with the additional snapshot tag, the following items are specifically updated on the LandingPage:

- The title (Figure 4-1)
- Date of publication (Figure 4-2)
- The requested snapshot is shown by default (Figure 4-3)
- The quote proposal (Figure 4-4)

Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) - Snapshot of Argo GDAC of May, 8th 2015

Publication date 2015-05-08

Author(s) Argo

DOI 10.17882/42182

Publisher SEANOE

Keyword(s) float, Argo, global ocean observing system, ocean circulation, in-situ, ocean pressure, sea water salinity, sea water temperature, multi-year, weather climate and seasonal observation, global-ocean, Installations de suivi environnemental

Abstract Argo is a global array of 3,000 free-drifting profiling floats that measures the temperature and salinity of the upper 2000 m of the ocean. This allows, for the first time, continuous monitoring of the temperature, salinity, and velocity of the upper ocean, with all data being relayed and made publicly available within hours after collection. The array provides 100,000 temperature/salinity profiles and velocity measurements per year distributed over the global oceans at an average of 3-degree spacing. Some floats provide additional bio-geo parameters such as oxygen or chlorophyll. All data collected by Argo floats are publicly available in near real-time via the Global Data Assembly Centers (GDACs) in Brest (France) and Monterey (California) after an automated quality control (QC), and in scientifically quality controlled form, delayed mode data, via the GDACs within six months of collection.

Licence

Utilisation A user of Argo data is expected to read and understand this manual and the documentation about the data contained in the "attributes" of the NetCDF data files, as these contain essential information about data quality and accuracy. A user should acknowledge use of Argo data in all publications and products where such data are used, preferably with the DOI and following standard sentence: "These data were collected and made freely available by the international Argo project and the national programs that contribute to it."

User's manual <http://dx.doi.org/10.13155/29825>

File	Size	Format	Processing	Access	Key
2015-05-08	6 GB	NC, NetCDF	Quality controlled data	Open access	42336

More snapshots

Data <ftp://ftp.ifremer.fr/ifremer/argo>
<ftp://usgodae.org/pub/outgoing/argo>

How to cite

Argo (2015). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) - Snapshot of Argo GDAC of May, 8th 2015. SEANOE. <http://doi.org/10.17882/42182#42336>

Click to download the data

Download metadata
TXT, RIS, XLS

Related publications

Fousiya T. S., Parekh Anant, Gnanaseelan C. (). Interannual variability of upper ocean stratification in Bay of Bengal: observational and modeling aspects. Theoretical and Applied Climatology, -

Kawakami Yuma, Sugimoto Shusaku, Suga Toshio (2016). Inter-annual zonal shift of the formation region of the lighter variety of the North Pacific Central Mode Water. Journal of Oceanography, 72(2), 225-234.

Related datasets

Feucher Charlene, Maze Guillaume, Mercier Herle (2016). 2000-2014 climatology of the North Atlantic permanent pycnocline properties. Seanoe.

Share

Figure 4 : LandingPage of a specific monthly Argo DOI. The #42336 identifies the monthly snapshot (<http://doi.org/10.17882/42182#42336>)

2.2. The DOI of French oceanographic cruises

2.2.1. Objectives for the cruises' DOIs

On behalf of the French scientific infrastructure for research fleet (TGIR), Sismer (the French NODC) assigns a DOI to each French oceanographic cruise. The main objective for a cruise DOI is to describe the cruise as a scientific activity and to improve the visibility of the Fleet TGIR as an infrastructure.

DOIs aim at building confidence and simplify the identification of publications that mention a cruise. If the authors cite the DOI in their scientific publications when they mention a cruise, it will be easy to spot these publications seeking the prefix DOI in full text publications. The number of publications is a key indicator of the efficiency of the fleet that TGIR has to provide its supervisory ministries.

These DOIs are an opportunity to centralize in a single page, information from many different systems to improve the visibility of all of this information. In this perspective, the cruises LandingPages list the datasets related to cruises published and archived in various data centers such as Sismer (the French NODC), Sedoo (a French scientific data center), Pangaea (an international ICSU data center) or Seanoe (a data publisher).

The cross references of DOIs between data centers improves an Internet the visibility and accessibility of cruise research data.

2.2.2. Organization

The cruise managers are required to post a Cruise Summary Report (CSR) at the end of the cruise to Sismer. Sismer⁴ manages the CSR and records the metadata of the cruise in a database. Then Sismer assigns a DOI through the DataCite API. The DOI LandingPage is built by merging the information gathered from multiple databases. The LandingPages are daily updated with a batch. The DOIs information is available in the catalog of French oceanographic cruises.

For the backlog of historic cruises, DOI has also been assigned to all historic French cruises referenced in the Sismer catalogue. Over 6000 such DOIs have been attributed to cruises. Returns on this initiative were overall very positive and authors now begin to quote the DOI of these cruises in their publications.

2.2.3. Series of cruises

Some oceanographic cruises are part of multiannual programs. A specific type of DOI has been attributed to series of cruises. A scientist can publish on and cite a series of cruises with the "series" DOI, or/and cite specific DOIs of cruises.

The LandingPage of the DOI on a series of cruises and its associated specific LandingPage cruise DOIs provides cross-links, useful to Internet users who will navigate between series and individual cruises.

2.2.4. Content of a sample Landing Page a cruise DOI

The LandingPage for a cruise DOI is designed to provide quick overview of all information related to the cruise. The cruise LandingPage is a "hub" of resources distributed different places. The information is presented in detailed blocks described in this paragraph. To allow a first synthetic vision of the cruise, by default, some blocks are partially hidden. Icons (Figure 5-12) are used to expand and collapse the blocks.

General data (Figure 5-1)

This is the main block; it presents the cruise as an experiment. The data in this block are mainly issued from the CSR sheet. In particular, we find:

- The title of the cruise
- The list of authors with possible links to their CV or ORCID record.
- A possible link to the cruise series if the cruise is part of it.
- The objective of the cruise
- ...

⁴ A project is underway to automate the after cruise completion metadata collection

Navigation map (Figure 5-2)

A map shows the navigation of the cruise. The samples are located on the map with green dots. The overview of these points displays information on the samples.

How to cite (Figure 5-3)

A block suggests the cruise citation as follows:

Nom des responsables de campagnes (Année de réalisation de la campagne) Nom de la campagne, RV Nom du navire, DOI

Links to external documents or sites (Figure 5-4)

The LandingPage is an information hub that provides links to documents or associated external sites (e.g. a cruise blog, a valuation record ...)

Videos of underwater vehicles (Figure 5-5)

The site provides links to the videos collected by underwater vehicles deployed during the cruise if these videos are freely accessible on the website videos Sciences Marines.

Published data (Figure 5-6)

This block lists published datasets related to the cruise. Datasets with a DOI are referenced in this section. If published in Seanoe, a link between the dataset and the cruise landing page is automatically set.

If a data set is published in another data warehouse (e.g. Zenodo, Pangaea ...), the author should consider to communicate the dataset DOI to Sismar. Sismar will then setup a cross-reference link.

Data managed by Sismar (Figure 5-7)

This block provides access to cruise data archived at Sismar. Most of these data come from automatic sensors and is delivered directly to Sismar. Some data are open access; others require the agreement of the cruise chief scientist.

Sampling operations (Figure 5-8)

This block lists all samples taken during the cruise. For each sample it provides a series of information.

Diving operations (Figure 5-9)

This block list the diving operations made by underwater vehicles during the cruise. It proposes, for each diving operation, a set of information and the diving trajectory. The course of the dive is illustrated with screenshots of devices available in the Marine Science Video web site.

Moorings (Figure 5-10)

This block details the moorings deployed during the cruise.

Bibliography (Figure 5-11)

This block lists publications where the cruise is cited. These publications are identified by a regular survey performed by the Library La Perouse. All referenced publications are registered, counted and processed in Archimer bibliography data base. This survey provides the fleet TGIR a set of bibliometric data on these publications. Using cruise DOIs in scientific publications should facilitate this work through reliable and easier tracking of publications.

French Oceanographic Cruises

MOMARSAT2014

1

Type: Oceanographic cruise
 Set: This cruise is part of the set MOMARSAT - MONITORING THE MID ATLANTIC RIDGE
 Ship: Pourquoi Pas?
 Dates: 13/07/2014 (Horta) - 31/07/2014 (Horta)
 Chief scientist(s): SARRADIN Pierre-Marie, CANNAT Mathilde
 DOI: 10.17600/14000300
 Objective: The cruise's objective was to perform the annual maintenance of the seafloor observatory set up on the Lucky Strike hydrothermal field during the MoMARSAT 2010 cruise. This observatory, which can transmit its data in near real time via acoustic, then Hertzian, channels, is one of the nodes in the EMSO (European Multidisciplinary Seafloor and water column Observatory) network. Its objective is to study over the long-term of active processes linked to seismicity, volcanism and hydrothermalism of the Mid-Atlantic Ridge and associated ecosystems. It includes both an infrastructure for data transmission (SEAMON stations on seafloor and BOREL surface buoy) and a set of sensors (geophysics, chemistry, physics, oceanography and biology). Some sensors are connected to the SEAMON stations and communicate all or part of their data towards shore every 6 hours, while the others are autonomous and store their data inside. As well as the observatory's technical maintenance, the second objective of MoMARSAT 2014 was to acquire measurements and spot observations on the Lucky Strike field in order to validate the time series data, extend the range of parameters studied and improve the spatial resolution of the study. This cruise is part of the EU EMSO European multidisciplinary seafloor and water column observatory project.

2

3

4

5

6

12

7

8

9

10

11

PUBLISHED DATA

Legrand Julien, Sarradin Pierre-Marie, Cannat Mathilde (2015). **Buoy management unit 1 data from the EMSO-Azores observatory, 2013-2014**. Seanoee. <http://dx.doi.org/10.17882/39765>

Legrand Julien, Sarradin Pierre-Marie, Cannat Mathilde (2015). **Buoy management unit 2 data from the EMSO-Azores observatory, 2013-2014**. Seanoee. <http://dx.doi.org/10.17882/39764>

Legrand Julien, Sarradin Pierre-Marie, Cannat Mathilde (2015). **COSTOF, technical data from SeaMoN West, node of the EMSO-Azores observatory, 2013-2014**. Seanoee. <http://dx.doi.org/10.17882/39764>

Legrand Julien, Sarradin Pierre-Marie, Cannat Mathilde (2015). **Iridium GPS 2 data from the EMSO-Azores observatory, 2013-2014**. Seanoee. <http://dx.doi.org/10.17882/39765>

Legrand Julien, Sarradin Pierre-Marie, Cannat Mathilde (2015). **Iridium GPS 1 data from the EMSO-Azores observatory, 2013-2014**. Seanoee. <http://dx.doi.org/10.17882/39766>

Legrand Julien, Sarradin Pierre-Marie, Cannat Mathilde (2015). **Meteorological data from the EMSO-Azores observatory, 2013-2014**. Seanoee. <http://dx.doi.org/10.17882/42583>

Sarradin Pierre-Marie, Cannat Mathilde, Blandin Jérôme, Ballu Valérie, Daniel Romuald, Godfroy Anne, Legrand Julien, Laks-Huon Agathe, Rommevaux-Jestin Céline, Sarrazin Joëlle, Colaco Ana, Coall Jean-Yves, Coulombier Thibaut, Guyader Gérard, Pichavant Pascal, Prigent Sébastien (2015). **EMSO-Azores seafloor observatory real-time data 2014 deployment**. SISMER. <http://dx.doi.org/10.12770/0240294-0346-4760-bc67-4401a583f475>

Sarradin Pierre-Marie, Legrand Julien, Cannat Mathilde (2015). **Turbidity (NTU) data from the EMSO-Azores seafloor observatory, 2014-2015**. SISMER. <http://dx.doi.org/10.12770/0240294-7667-474d-a115-99f0a67568b>

Sarradin Pierre-Marie, Legrand Julien, Cannat Mathilde (2015). **Temperature (°C) and dissolved oxygen concentrations (µmol/l) data from the EMSO-Azores seafloor observatory, 2014-2015**. SISMER. <http://dx.doi.org/10.12770/0240294-60f4-44b1-9c66-138a86685d33>

DATA MANAGED BY SISMER

SAMPLING OPERATIONS

DIVES

MOORINGS

BIBLIOGRAPHY

Marabos Marjolaine, Cuvelier D., Brouard J., Shilto B., Ravoux J., Zindeni M., Barthélemy D., Sarradin Pierre-Marie, Sarrazin Joëlle (2015). **Behavioural study of two Hydrothermal crustacean decapods: Mirocaris fortunata and Segonzacia mesatlantica, from the lucky strike vent field (mid-Atlantic ridge)**. *Deep-sea Research Part II-topical Studies In Oceanography*, 121, 146-158. Publisher's official version : <http://dx.doi.org/10.1016/j.dsr2.2015.04.008> , Open Access version : <http://archimer.ifremer.fr/doc/00266/37714/>

Sarrazin Joëlle, Legendre Pierre, De Bussierolles Fanny, Fabri Marie-Claire, Gullini Katja, Ivanenko Viatcheslav N., Morineaux Marie, Vanreusel Ann, Sarradin Pierre-Marie (2015). **Biodiversity patterns, environmental drivers and indicator species on a High-temperature hydrothermal edifice, mid-Atlantic ridge**. *Deep-sea Research Part II-topical Studies In Oceanography*, 121, 177-192. Publisher's official version : <http://dx.doi.org/10.1016/j.dsr2.2015.04.013> , Open Access version : <http://archimer.ifremer.fr/doc/00266/37713/>

Laes Agathe, Legrand Julien, Sarradin Pierre-Marie, Rolin Jean-Francois, Tanguy Virginie, Cathalot Cecile, Blandin Jerome (2015). **Long term in situ monitoring of total dissolved iron concentrations on the MoMAR observatory**. EGU2015 - European Geosciences Union General Assembly, 12 - 17 April 2015, Vienna.

Opderbecke Jan, Brisset Laurent, Boy Emillie (2015). **Infrastructures Marines et Numériques. Unité Systèmes Sous-Marins. Rapport d'activité 2014**

Laes-Huon Agathe (2014). **Compte Rendu de Réunion. CHEMINI - AISO802**

Laes Agathe, Legrand Julien, Tanguy Virginie, Cathalot Cecile, Blandin Jerome, Rolin Jean-Francois, Sarradin Pierre-Marie (2014). **Monitoring à long terme des concentrations en fer total dissous sur l'Observatoire MoMAR : Long term in situ survey of total dissolved iron concentrations on MoMAR observatory**. *Sea Tech Week 2014 - 9th Marine Science Conference - From 13th to 17th October 2014 - Brest, France*

Figure 5 : DOI LandingPage of Momarsat 2014 cruise <http://dx.doi.org/10.17600/14000300>