Deliverable

Data Management Plan

LIFEDEEPER: LIving together in the Future: vulnErability of DEEP sea Ecosystems facing potential mineral Resources exploitation

WP 1
Version : 1
Creation date : 2022-11-18
Last modification date : 2023_03_03
License : CC-BY
Contributors : Cotten Clémence (Ifremer), Cambon Marie Anne (Ifremer), Constance Pagès (Ifremer), Valérie Harscoat (Ifremer)

This research was funded through the French Gouvernment managed by l'Agence Nationale de la Recherche au titre de France 2030 : **ANR-22-POCE-0007**

LIFEDEEPER LIving together in the Future: vulnErability of DEEP sea Ecosystems facing potential mineral Resources exploitation

The protection of the deep seabed, the largest biome on the planet, is a major challenge for future generations. Covering more than 70% of the Earth's surface, the ocean encompasses more than 80% of the volume of seawater, produces at least 50% of the atmospheric oxygen and absorbs ≈25% of anthropogenic CO_2. Hidden from our eyes, the deep ocean, aphotic, under pressure, has long been considered as a vast desert of darkness, sometimes populated by sea monsters. However, today it is the covetousness of developed and emerging countries for its resources in terms of biodiversity and minerals, particularly in hydrothermal contexts. Within the framework of structuring programs for sustainable development (e.g. the UN *Ocean Decade*, calls for the climate (COP26), *France 2030*), it becomes urgent to acquire, through holistic approaches, solid knowledge baselines with regard to the diversity of geological systems, habitats, taxonomic and functional biodiversity, at open and interconnected ocean scales.

Our consortium leads a five cruises series (2014- 2023/4) in the framework of the French contract for exploration of Atlantic polymetallic sulphides resources (ISA). At 3600 meters depth, the TAG hydrothermal field is an area of choice for studying the geo-biodiversity of so-called inactive SMS deposits, and is, with the young and volcanic Snake Pit hydrothermal field, at the heart of our studies. Separated by 300 kms of ocean ridge and a major transform fault shifting its axis by 150 km, these sites show similarities in terms of biological diversity, raising the question of their connectivity. Exploration of the area revealed geochemical evidences for unknown active vent sites located between TAG and Snake Pit and further south. These sites constitute potential relays for the organism dispersal, a key point in the overall understanding of ecosystems and their relationships along ridge segments.

LIFEDEEPER intend to develop new approaches, combining *in situ, in vivo*, and lab experimentations, modeling and qualitative research in social sciences to disentangle the geological, geochemical and biological natural functioning of deep ocean ecosystems. In addition to the coordination work package (**WP1**), 4 multidisciplinary, complementary scientific WPs and one communication WP are proposed:

WP2: Exploration and description of ecosystems associated with inactive and active massive sulphide deposits, toward integrated definitions of active and inactive vents.

WP3: Integrated 3D study coupling hydrodynamics, distribution of trace metals and numerical modeling to assess the biogeochemical impact of the hydrothermal plume to the deep ocean.

WP4: Study the connectivity and life cycle of holobiont models, capacity of adaptation and acclimation allowing the resilience of communities.

WP5: Legal and political analysis of international regulatory regimes, sociological analysis of science-technology-society, capacity building.

WP6 will then aim to produce effective educational and public outreach activities targeting citizens, students, scholars, organizations and various stakeholders, through participative science, educational science and art.

Ultimately, we want to establish a precise map of contrasting sites at the scale of a 600 km ridge segment, in order to provide key parameters to understand the natural functioning of these environments, both from a geological and biological point of view. Key components of the functioning of these ecosystems and associated services, along activity gradients, will permit establishing global-scale inter-comparison protocols. In a holistic and multidisciplinary approach, *LIFEDEEPER* aims to acquire definitions of reference ecological profiles and preservation strategies in the current context of the growing interest in deep mineral resources for the carbon-free and digitized world economy. *LIFEDEEPER* will propose solutions to allow future informed guidelines and decision-making.

Start date: 2022-09-21
End date: 2028-09-20

Funding

France 2030: ANR-22-POCE-0007
Ifremer
CNRS
UBO
Sorbonne Université
MNHN
IRD

Partners

- **Ifremer**
- **CNRS**
- **UBO**
- **Sorbonne Université**
- **MNHN**
- **IRD**

Table of contents

1. Purpose of this document

A Data Management Plan (DMP) is a formal document that describes the data management life cycle for the data to be collected, processed and/or generated by a research project (European Commission, 2016). This evolving DMP will ensure that the data is well documented, archived, interoperable and accessible for reuse. Accessibility of the data will be made possible in part by the LIFEDEEPER website *(D1.1)* and the data catalogue *(D1.3).*

2. Data Management

Data management activities regarding scientific data collected within the project will be supported by SISMER (Scientific Information Systems for the Sea) in Ifremer (https://data.ifremer.fr/), for the data managed in its missions as National Oceanographic Data Centre (NODC) for France for the International Oceanographic Data Exchange program of UNESCO (IODE), and as one of the partner centres of the french ocean cluster ODATIS (https://www.odatis-ocean.fr/en/) in the Data Terra research infrastructure (https://www.data-terra.org/en/). Furthermore, SISMER handles since 2011 an ISO 9001 certification for its "Collect and make available data on the marine environment" process, and since 2019 a "Core Trustworthy Data Repositories" from the CoreTrustSeal certification of the RDA (Research Data Alliance).

3. Data identification & description

Raw data, physical objects and audiovisual content such as videos and photos will be collected during cruises planned throughout the project. Some data such as isotopic data or raw high throughput sequences will be produced with the materials collected during cruises (sediment cores, sub-sampling of faunal samples …).

Forthcoming cruises to collect sample for LIFEDEEPER experiments are:

- BICOSE3, scheduled 20th October to 5th December 2023
- A cruise dedicated to exploration of the MAR, lead D Jollivet/F Lallier, CNRS Roscoff. A proposal to explore a portion of ridge (34°N/S-OH2 to 6°N/Markov Deep) and revisite some specific vent sites (Broken spur, Logatchev) surrounding the SnakePit/TAG area will be submitted to the CNFH in September 2023.

The project will use raw data and audiovisual content collected from past cruises. These cruises are:

- BICOSE2014: https://doi.org/10.17600/14000100
- HERMINE 2017: https://doi.org/10.17600/17000200
- BICOSE2 2018: http://dx.doi.org/10.17600/18000004
- HERMINE2 2022: https://doi.org/10.17600/18001851

- BICOSE3 for end 2023 (20th of October till 5th of December).

All cruise reports will be available to the consortium.

Most of the data (faunal abundance, environmental and isotopic data) are in csv format, easy to re-use with many tools. Some of the data are in specific format, standard for the domain. For example, High Throughput Sequence are in FASTQ format and bathymetry in NetCDF format.

DATASETS

➢ High Throughput Sequences (DNA)

DNA sequences will be produced by using sediment cores collected by Human operated vehicle (HOV-Nautile – HERMINE2 and BICOSE3) or and multicorer (MUC) and USNEL in 2022 and 2023. The data will be available in 2024. It will be used by WP 2 and WP4.

DNA sequences will be produced by using rocks and water samples collected by Human operated vehicle (HOV-Nautile) in 2022 and 2023 (both sequencing crude samples and/or after enrichment cultures –batch or continuous such as gas lift fermentation). The data will be available in 2024. It will be used by WP 2 and WP4.

DNA sequences will be produced from animal samples for identification (WP2), connectivity, genome approaches (WP4).
The project will also re-use DNA sequences that already exist in the public database Sequence Read Archive (project code: PRJNA544999). The data will be Open Access within 4 years of submission or upon publication, whichever comes first. The format of these data is FASTQ files[1], which is standard for the domain.

[1] FASTQ format is a text-based format for storing both a biological sequence (usually nucleotide sequence) and its corresponding quality scores (source: Wikipedia).

Also, DNA sequences will be produced by analyzing colonization on substrates during HERMINE2 2022 along the activity gradient of the TAG district, and that will be recovered during BICOSE3 asked in end 2023. All data will be deposited in public databases.

- **Faunal data**

Faunal data includes abundance, density, diversity, morphometric measurements, fecundity and biomass datasets. This data will be acquired by processing faunal assemblages and organisms collected with a submersible (arm grab, suction sampling, coring and substratum deployments). Faunal data from project cruises and past cruises will be used for WP 2 and 4.

Faunal data from past cruises are available in open data on the repository Seanoe and on other partners" repositories; e.g.: Pangea.

The format of all these data is csv. The taxonomic repository used is the World Register of Marine Species (WoRMS).

- **Environmental data**

Environmental data includes temperature, O_2, CH_4, H_2S, CO_2, metal concentrations, currents, turbidity, salinity, metabolic rate measurements, microbial cell concentrations and organic matter concentrations measurements. These data are collected using a submersible through (*in situ* measurements (analyzer, temperature probes, microprofiler and autonomous probes) or through fluid/water collection. Environmental data from project cruises and past cruises will be used for WP 2, 3 and 4.

Environmental data from past cruises are available in open data on the repository Seanoe on another partners repositories; e.g.: Pangea.

Geochemical data, including major and trace element composition of rocks, sediments and fluids (WP3) will be deposited on seanoe with a DOI assigned and made publicly available upon publication.

The format of all these data is csv.

- **Isotopic data**

Isotopic data includes carbon (^{13}C and ^{14}C), nitrogen, copper, iron, zinc, germanium, selenium, silicium, strontium and sulfur measurements. These data are produced by sub-sampling faunal, fluid, sediment, and rock samples in the lab at Ifremer and other research institutes. These samples will be collected during project cruises, but isotopic data from samples collected during past cruises will be used as well.
Faunal isotopic data from past cruises in the deep sea are available in open data on the Seanoe repository (https://doi.org/10.17882/76595, https://doi.org/10.17882/84195).

The format of this data is csv.

Isotopic data from past cruises are currently submitted for publication and will be made available in open data on the Seanoe repository. New isotopic data generated during the project will be also deposited on seanoe with a DOI assigned and made publicly available upon publication.

- **Video imagery**

Most partners collect video imagery from submersible or towed camera during cruises. Video sequences from project cruises and past cruises will be used for WP 2, 3, 4, 5 and 6.
Video imagery from past cruises are available in open access on SISMER video database "Vidéos Sciences Marines" (http://video.ifremer.fr/). The format of these videos is mp4.

- **Photos**

Most partners collect photos during each cruise. Photos from cruises of the project and photos from past cruises will be used by all work-packages.
Ifremer photos of past cruises are available on open access on the Ifremer's Oceanothèque.
The format of photos is jpeg.

- **2D/3D mosaics and models**

Images and videos will in certain case be used to reconstruct 2D/3D mosaics or 3D point models from photogrammetry technics. Reconstructions format is .tif, .geotif, .mnt, .obj. Reconstructions are available in open access in Seanoe (https://doi.org/10.17882/79218).

- **Bathymetry**

Bathymetry datasets are collected by multibeam sonars from the oceanographic vessels by most partners. Bathymetry datasets from project cruises as well as bathymetry from past cruises will be used in WPs 2-3 and 4. Raw bathymetry from past cruises of Ifremer are stored on the SISMER geoscience database and accessible by Ifremer campaign catalog (bathymetry datasets are available on each cruise description) and by SeaDataNet Common data index (CDI). Digital Terrain Models (DTM) are stored and accessible by Sextant.
The format of bathymetry datasets is NetCDF which is common for the marine domain, and Digital Terrain Models is TIF.

- **Geospatial data**

Maps will be included in publications to locate operations, samples and transects. In addition, large-scale maps will be created for the WP5.

- **3-D coupled hydrodynamical-biogeochemical numerical modelling**

We use the Coastal and Regional Ocean Community Model (CROCO, https://www.croco-ocean.org) whose code is freely available online (https://www.croco-ocean.org/download-2/) along with the biogeochemical model PISCES (https://www.pisces-community.org) freely available through different distributions (https://www.pisces-community.org/index.php/code-distribution/). The data we generate are 3-D model outputs of velocity, temperature, salinity and geochemical compounds in NetCDF format (https://www.unidata.ucar.edu/software/netcdf/). They are stored on Datarmor and can be made available upon request. However, no DOI is traditionally assigned to regional model outputs such as the ones that we generate.

- **Cruise metadata**

Cruise metadata collected during each cruise as well as metadata from past cruises will be used by WP 2, 3, 4, 5, 6.
The metadata of Ifremer cruises is freely available on the French Oceanographic cruises directory (http://campagnes.flotteoceanographique.fr/). All metadata of past cruises used for the project are freely available on the Cruise Summary Report of Seadatanet (https://csr.seadatanet.org/)
The format of cruise metadata is the CSR[2] (Cruise summary reports) format, which is used at European level by all organizations involved in the SeaDataNet project. It ensures the interoperability of cruise metadata between all project partners. Cruise metadata is available on average one year after the cruise occurred.

- **Interview recordings and transcriptions**

Interview recordings and transcriptions will be collected and used during the project by WP5. Depending of the type of interview (directive, semi-directive) and their organization, interviews will be recorded in sound or video format, and transcribed in text or Excel files *(see section 7*).

- **Questionnaire and survey data**

Questionnaire and survey data will be collected and used during the project by WP5. They will be organized in excel files.

[2] To know more about this format : https://www.seadatanet.org/Standards/Metadata-formats/CSR

PHYSICAL OBJECTS (SAMPLES)

- **Sediment/rock, faunal and microbial collections**

Most partners collect sediment/rock, faunal and microbial materials during each cruise. These materials will be used for WPs 2, 3 and 4. Physical objects will be kept in the different laboratories. Faunal and microbial collections will be preserved in ethanol, formaldehyde (room temperature), at 4°C or frozen (-20°C and -80°C). The material not used for analyses could be accessible on demand to the PIs.

DOCUMENTS

- **Reports, publications, posters, meeting and workshop reports, communication material and educational resources**

Documents will be produced throughout the project by all project's partners from WP 1 to WP 6.
The format of documents is pdf which ensures that the documents can easily be consulted by all, and if necessary documents will also be published on .doc and .odt format for easier reuse.

AUDIOVISUAL

- **Video clips**

Ifremer will produce video clips in 2023 for WP6. The format of these video clips will be mp4.

- **Photos**

During events, photos and/or videos will be taken to contribute to the communication objectives of LIFEDEEPER.
The format of photos is jpeg.

- **Website**

Ifremer created the LIFEDEEPER website (https://lifedeeper.ifremer.fr) in 2023 for WP 1 using eZ CMS. Ifremer will be in charge of updates and management of the website throughout the project.

- **Newsletter**

A newsletter will be created and send to all participants of the project two times a year by Ifremer for WP 1, amended by all WP leaders.

- **Mailing list of potential research participants**

Ifremer maintains several mailing lists of all LIFEDEEPER participants including one dedicated to the institutional PIs (lifedeeper@listes.ifremer.fr), one dedicated to WP leaders (lifedeeperwpleader@listes.ifremer.fr), one dedicated to the entire LIFEDEEPER community (lifedeeperall@listes.ifremer.fr) and one dedicated to the advisory board (advisory_board_LIFEDEEPER@listes.ifremer.fr). A generic email address (lifedeeper2022@ifremer.fr) is used to communicate with the coordinating board (MA Cambon, E Pelleter and AL Le Velly)

UBO-Amure maintains a mailing list with the name, gender, email and organizational affiliation of other participants (i.e., stakeholders and researchers interested in deep-sea mining issues or the conservation of deep-sea ecosystems) to research activities (workshops, focus groups, interviews, surveys) undertaken in the WP5. These personal data are organized in an Excel file following RGPD requirements (see part 7).

4. Data organisation & exchange (internally, during the project)

The data manager will ensure the good application of the DMP and will coordinate the work of the data collectors/providers/user from each WP. WP leaders and other project's participants have to inform the data manager about (1) which datasets will be used, collected or created throughout the project, (2) where they will be stored and (3) when and how they will be published in order to update the DMP. The data manager will help project's participants to store and publish their datasets and other materials, choose a license, create a DOI, etc. As much as possible, it is recommended to organize folders by cruise, work-packages or deliverables. The file names should be as complete as possible, with a date and/or a version following this model: DeliverableNumber_Name_Version_DateYYMMDD

In each work-package, researchers from different organizations will have to work together and exchange datasets. In order to facilitate daily exchanges and avoid the extra energy consumption associated with sending attachments by email, we recommend the use of a private cloud. Ifremer will provide **Owncloud** spaces for project participants. The difference with a public cloud is that in Owncloud, data is stored within the institute's infrastructure and access is subject to the institute IT security policy. At Ifremer, access to Owncloud is limited to people with an extranet account and authorization to access a document will be given individually. Interview recordings, transcriptions, questionnaire, survey data and mailing lists of potential research participants will be stored on Ifremer's Owncloud. The space available on the cloud for each user is limited to 5 Gb. If data sets are too large, project's members can use a web-based application for sending large files, like Filesender (a Renater service).
For long storage of data and achieving, an **Alfresco** spaces dedicated to LIFEDEEPER will be created, accessible through the same extranet account.

The majority of datasets will be used by several WPs requiring to make some data sets available to all project participants. The data catalogue on the LIFEDEEPER website will provide access to datasets to all

project participants through restricted access (in the case of period of exclusivity for example). Each dataset will be associated with a metadata sheet in order to facilitate their usability.

5. Data archiving: data and metadata storages and repositories

Data archiving and publications will be based mainly on repositories, database and other services operated by SISMER (Scientific Information Systems for the Sea) which is certified Core Trustworthy Data Repositories by the Research Data Alliance. Publications will also be based on HAL repository.

- **Datarmor supercomputer (Ifremer)**

If necessary, for large data sets, Ifremer will propose a data storage space on the supercomputer Datarmor. This storage space ensures the durability of the data stored.

- **SISMER geoscience database (Ifremer-SISMER)**

Bathymetry datasets collected by multibeam sonar from French oceanographic fleet ships are permanently archived on magnetic tapes and indexed on the Sismer geoscience database. After processing, they will be available approximately one year after the cruise. These data are stored on the native format of the instrument and on NetCDF format. Basic quality controls are carried out by the automated system of the database (coordinates, dates …). The Ifremer Marine Geosciences research unit's team use the software "Globe" to validate the bathymetry data and produce a qualified dataset, an intermediate dataset between raw bathymetry data and a Digital Elevation Model (DEM).
Bathymetry datasets, stored on the SISMER geoscience database, can be found on the French Oceanographic cruises directory [3] on the metadata description sheet of each cruise, and on the SeaDataNet Common Data Index (CDI)[4]. They are also used by the Emodnet bathymetry project to create DEM on open data.

- **BIGOOD - BIology and GeOlogy Ocean Database (Ifremer-SISMER)**

Faunal data (abundance and diversity) collected from the French oceanographic fleet ships are archived on the BIGOOD database. BIGOOD is part of the cruise database managed by Sismer (which also includes the geoscience database and cruise metadata) and includes metadata and data related to marine geological and biological samples collected onboard by IFREMER teams and/or stored in its repository in Brest (CREAM; sampling operations (location, sampling equipment), descriptions of samples, samples storage (physical location and movements), related analysis (faunal data (abundance and diversity), chemistry, granulometry, stratigraphy, lithology, etc.). The sample metadata are accessible on the French

[3] https://campagnes.flotteoceanographique.fr/
[4] https://cdi.seadatanet.org/search

Oceanographic cruises directory (https://campagnes.flotteoceanographique.fr/) and on the Geological and biological marine sample collection website (https://echantillons.ifremer.fr/en). For each geological sample stored in its repository, IFREMER allocates an International Generic Sample Number (IGSN) a persistent international unique identifier for physical samples and specimens. IFREMER is an IGSN allocating agent (https://www.igsn.org/allocating-agents/) and ensure persistent access to the metadata record of the registered IGSN.
BIGOOD database guarantees data standardization and the use of standardized vocabularies such as OBIS for taxonomy and P01 (BODC) for parameters. BIGOOD's datasets are exported to the Ocean Biodiversity Information System[5] (OBIS), which data is integrated into the Global Biodiversity Information Facility[6] (GBIF) thanks to a collaboration agreement[7] between the two institutions. A unique identifier identifies each sampling. BIGOOD is part of Ifremer's ISO 9001 quality system. A quality control protocol was therefore developed. However, BIGOOD does not offer a data dissemination service, exports must be ensured through a repository such as Seanoe.

- **Seanoe (Ifremer-SISMER)**

Faunal, environmental and isotopic data will be stored on the repository Seanoe (Sea Scientific Open data edition). After processing, data will be available for an average of two years after the cruise. Seanoe is a scientific marine environment data publication solution created by the French NODC (National Oceanographic Data Centre). It allows scientists to publish a dataset with open access and to reliably cite it in an article. Total file size limit per record is 100GB but it is possible to publish multiple datasets. The long-term preservation of data filed in Seanoe is ensured by Ifremer infrastructure. Data published by Seanoe are free to use. It is described by metadata information, which can be converted on OIA-PMH (open archive initiative – protocol for metadata harvesting), Data cite or ISO 115 formats to ensure their interoperability with other tools. All metadata information is checked by SISMER's team before publication. *Link: https://www.seanoe.org/*

- **SISMER video database "Vidéos Sciences Marines" (Ifremer-SISMER)**

Video imagery acquired by the French oceanographic fleet will be stored on the Sismer video database under "scientific underwater images". Video sequences are organized by oceanographic cruise, submersible, dive and camera type. Metadata relative to submersible dives and acquisition conditions are also provided, they include: submersible position and altitude, temperature and salinity conditions, camera position and configuration. Comments recorded simultaneously with the videos are displayed if present. After processing the data, it will be available for an average of one year after the cruise.
Link : http://video.ifremer.fr/index

[5] https://obis.org/
[6] https://www.gbif.org/fr/
[7] https://obis.org/2014/10/03/gbif/

- **Oceanotheque (Ifremer-SISMER)**

Photos acquired by the French oceanographic fleet will be stored on the Ifremer's "Océanothèque" which is the audiovisual archive of still and moving images.
This project is part of the international Open Access principle which seeks to make scientific documentation accessible to the greatest number of people by broadcasting it free of charge on the web.
Link : https://image.ifremer.fr/search

- **Archimer (Ifremer-SISMER)**

Publications, reports and documents like posters will be stored on the Ifremer's institutional repository Archimer. It contains all the documentation produced by scientists employed by Ifremer. The scope of Archimer is also extended to papers mentioning French cruises as well as those produced by the UMRs with which Ifremer is associated.
The long-term preservation of documents filed in Archimer is ensured by Ifremer infrastructure. They are open access and described by metadata information, which can be converted to the Dublin-core format OIA-PMH (open archive initiative – protocol for metadata harvesting) to ensure their interoperability with other tools. Archimer offers the possibility to create links with related datasets or cruises. The deposits in Archimer are controlled and validated by the Ifremer library team.
Link : https://archimer.ifremer.fr/search

- **HAL (University, CNRS)**

The French open archive HAL was developed by the French Center for Direct Scientific Communication (CCSD) for the deposit and dissemination of scientific articles (published or not), theses and other objects. This multidisciplinary archive allows the dissemination of knowledge of the entire French scientific and academic community and the construction of many services (HAL website).

The CCSD is under the joint supervision of the National institute for Research (CNRS), the National Institute for Research in Digital Science and Technology (INRIA) and the National Research Institute for Agriculture, Food and Environment (INRAE). HAL is hosted on the CNRS' own data centers in Villeurbanne (France), ensuring long lasting and trustworthy access to data.

Some participants to the project are expected by their institutions to upload their publications, reports and documents on HAL, even if these have been posted to Archimer as well.

- **Sequence Read Archive**

The Sequence Read Archive (SRA) of NCBI (National Centre for Biotechnology Information) which is part of the NIH (National Institutes of Health) funded by the USA government will be used to stored High Throughput Sequence (DNA) in zipped FASTQ format. The SRA data, available through multiple cloud providers and NCBI servers, is the largest publicly available repository of high throughput sequencing data. The long-term preservation of HTS data SRA is ensured by NIH infrastructure. Each dataset receives a unique identifier that can be searched for in the SRA website to gain access to the raw data as well as the linked metadata sheet and attributes file. Data can be made public immediately after processing, at user-specified chosen date, or after publication. The submissions in SRA are controlled and validated by the SRA team.
Link: https://www.ncbi.nlm.nih.gov/sra.

6. Data publishing & licensing

All data produced during the project has to be easily accessible. The majority of tools used to store data have a portal where users can easily find datasets in opendata. Seanoe, the Ifremer's "Océanothèque" and Archimer contributes to Open Access / Open Science principle for free access to all publicly-funded scientific data for the benefit of research. Seanoe is developed within the framework of the ODATIS Ocean Cluster, which promotes the FAIR principles. Seanoe datasets can easily be reused by other projects. For example, data published in Seanoe are also automatically duplicated to the EMODnet Data Ingestion portal so that marine data centers of the EMODnet Ingestion network will be informed about the existence and availability of datasets published in Seanoe. Those data centers will then undertake activities for elaborating submitted datasets for uptake in their national databases and wider publishing in European data portals such as SeaDataNet, EurOBIS, and EMODnet thematic portals, if the datasets are of interest for their portfolio.

As a mining permit contractor, Ifremer is obligated to send its data to the International Seabed Authority[8] (ISA). Therefore, data acquired by Ifremer during the project will be transmitted to ISA.

In order to facilitate the findability of data and other materials of the project, two tools will be used :

- **LIFEDEEPER Website**

The LIFEDEEPER website is the centralized access point to all data and resources used and produced during the project. All reports, publications, posters, newsletter, communication material educational resources and meeting reports will be downloadable directly on the website. A "maps and data" section will provide

[8] https://data.isa.org.jm/isa/map/

links to the portals where the project data can be found and a data catalogue will be directly integrated on this section.
Link : https://LIFEDEEPER.ifremer.fr/

- **Data catalogue**

All project outputs (data, maps, models ...) will be managed according to FAIR principles into a data catalogue (D1.3) created and managed on the Spatial Data Infrastructure for marine data, operated by SISMER, called Sextant. Sextant follows the principles of FAIRness. It provides interoperable web services for discovery, viewing, and downloading data and products. These services use standards defined by the Open Geospatial Consortium and fulfill the INSPIRE Directive requirements.
Link : https://sextant.ifremer.fr/eng

LICENSES

As far as possible, data and other materials are to be published on CC0 or CC-BY license. Seanoe, Sextant and the Ifremer's "Océanothèque" allow data producers to choose which license to apply on their data. These tools also offer a way to cite data.

On the Sismer geoscience database, all bathymetry datasets are on CC-BY license (except data acquired in foreign waters).

While videos in the SISMER video database are not currently under a license, they are protected by intellectual property right. A written permission from Ifremer is requested before copying, translating, reproducing, selling, publishing, exploiting or disseminating a video. Video sequences can be downloaded at low and medium definition immediately, but for high definition, a request has to be sent to the chief scientist of the related cruise.

EMBARGO AND RESTRICTIONS TO DATA SHARING

The majority of data produced during the project (e.g. faunal, environmental and isotopic data) will not be accessible until the article(s) using the data is published. On Seanoe, an embargo of up to 2 years may be applied to a dataset (for scientific publications in progress). However, during this period of exclusivity, it will still be possible to store and describe datasets on Sextant, Seanoe or Archimer.

Indeed, these tools allow the administration of access rights in order to limit accessibility of some datasets to the project team only. Bathymetry datasets acquired in non-French national waters an stored in the

Sismer geoscience database, are not on open data. A request must be sent to the database administrator. Commercial reuse is not allowed.

CITATION AND ACKNOWLEDGEMENTS

Acknowledgements has to be included in publications, conferences and all project outputs:

This research was funded by the France 2030 French Government program, managed by the National Research Agency: ***ANR-22-POCE-0007***

Ce travail a bénéficié d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre de France 2030 : ***ANR-22-POCE-0007***

DOI

Each data set published in Seanoe will be assigned a DOI (Digital Object Identifier used as a resource identification mechanism) in which the author (data producer) is clearly identified. This allows data to be cited in an accurate, reliable and long-lasting manner. Sextant and Archimer also offer on-demand DOI creation. All three offer a fast responding service: a well described dataset can be assigned a DOI within 24 hours.

FAIR PRINCIPLES

Data management in LIFEDEEPER project will apply the FAIR principles. In order to be easily findable, datasets have to be described by metadata on each repository and identified by a persistent identifier. On the data catalogue, data will be described by a metadata sheet which respect a rich and interoperable format (ISO 19115) to be accessible.

7. Personal data

DATA TYPES THAT INVOLVE PERSONAL DATA

Interviews include personal data such as names of interviewees, their personal background and organizational affiliations that are required to understand the full context of the statements made and knowing whom to contact in case of further questions. This information is available only to LIFEDEEPER researchers and will be treated with confidentiality within the project. In publicly available products such

as reports and articles, anonymity will be the 'default' procedure for all interview data, unless otherwise agreed (e.g., if the interviewee has the opportunity to claim credit for the knowledge they offer to LIFEDEEPER).

Questionnaire and survey data will contain anonymized IDs of respondents so that the data entries cannot be linked to their personal identity. Questionnaires and survey data will be stored in Excel files on a secured Ifremer drive (Owncloud), where they will remain for the duration of the project's needs. This information is available only to LIFEDEEPER researchers and will be treated with confidentiality within the project.

UBO-Amure maintains a **mailing list** with the name, gender, email and organizational affiliation of potential participants (i.e., stakeholders and researchers interested in deep-sea mining issues or the conservation of deep-sea ecosystems) to research activities (workshops, focus groups, interviews, surveys) undertaken in LIFEDEEPER WP5. These personal data are organized in an Excel file and stored on a secured Ifremer drive (Owncloud), where it will remain for the duration of the project's needs.

During events, **photos and/or videos** will be taken to contribute to the communication objectives of LIFEDEEPER. This will be made clear to participants in advance so that they can choose not to be included in photographic material.

PERSONAL DATA MANAGEMENT

The collection and management of the data types involving personal data listed above shall comply with the GDPR principles[9].

The collection and processing of personal data for the interviews and surveys for social science research will fully comply with Regulation (EU) 2016/679 (General Data Protection Regulation) regarding the collection and processing of personal data and the protection of privacy.

The researchers will obtain, prior to any data collection, informed consent from all participants in the interviews and surveys. If requested by the data security regulations, the data will be handled and stored in a non - identifiable (anonymous) format, i.e. the data entries cannot be linked to the personal identity of the study participants.

The collection of personal data as part of LIFEDEEPER will be undertaken mainly through surveys and questionnaires circulated to volunteers signing up and participating in workshops or focus groups. It should be noted that the collection of personal data per se is not the purpose of LIFEDEEPER, but certain

[9] Complete guide to GDPR compliance : https://gdpr.eu/

aspects of the data may be necessary for research (e.g. information on the member State/region in which an individual is located).

LIFEDEEPER partners are long-established organizations that are involved in many research projects. They thus fully comply with the applicable EU and national legislations on the issue of personal data protection; some also have internal procedures in place on these aspects. If the organizations have their own ethics committees, they will be consulted. Data transfer will only take place among LIFEDEEPER partners by email or via a secured access online platform.

Whenever products are derived from these data, the data will be pooled and anonymized, such that individual respondents cannot be identified. All external participants contributing to the research will be provided with a detailed information sheet on the nature of the work being conducted to ensure their complete understanding. The information sheet specifies the aims and methods of the study. It also reminds participants that their involvement is voluntary and that they can withdraw their participation at any time. For participants who require it, translation of the information sheets will be provided.

INFORMED CONSENT PROCEDURES FOR INTERVIEWEE

Informed consent is a key principle of ethical research. By adhering to this principle, LIFEDEEPER will ensure that research participants are involved in the project entirely on a voluntary basis and that they are adequately informed of the potential risks and benefits of their participation. Participants will also be informed that all data gathered about them remains under their control.

The key ethical principles of the project are based on:

- Participants must be informed fully about the purpose, methods and possible uses of their participation, what it entails and what risks, if any, are involved.

- Participants must participate in a voluntary way, free from any coercion and any decision not to participate must not lead to any negative consequences or the impression thereof.

When the researcher first contacts the interviewees (by phone, e-mail or in person), the researcher will inform them about the purpose of the interview (in written or oral form). At the beginning of the interview, the interviewer will re-cap the purpose of the interview; the storage and usage of the interview data; agree on anonymisation or credit/citation procedure; ask for permission to record the interview; and inform the interview participant about his/her right to withdraw his/her participation at any time during the research project.

The printed consent form (Annex 1) will be presented to the interviewee and, upon acceptance, signed. Preferably, consent forms will be sent to the interview participants prior to the interview, but this may

not be the case when interview opportunities develop spontaneously e g. on the same day per telephone. The signed forms are then securely stored by the LIFEDEEPER partner performing the interview. They can be kept in an electronic form if this is in accordance with national and European rules regarding electronic archives. Alternatively, consent can be given and recorded at the beginning of the interview.

The participant will be provided with a copy of the consent form, which includes contact information of the researcher as well as the DPO at the coordinator's institution (Jean-Marc Sinquin, dpo@ifremer.fr, Ifremer).

8. Re-use of existing data

Table 1 describes the main datasets that will be re-used in LIFEDEEPER. This data source will be compiled in the data catalogue. Of note, other sequences (genomes, transcriptomes, metagenomes, metacarcoding) and data obtained from other funded projects will be used in LIFEDEEPER.

Table 1 : Overview of re-use of data per WP

WP/task	Datasets	Authors	Data source	Are data open access?
WP 1	BICOSE and HERMINE cruises data	Cambon, Fouquet, Pelleter, Cathalot	BICOSE2014 : https://doi.org/10.17600/14000100 HERMINE 2017 : https://doi.org/10.17600/17000200 BICOSE2 2018 : http://dx.doi.org/10.17600/18000004 HERMINE2 2022 : https://doi.org/10.17600/18001851	Yes
WP 1	BICOSE and HERMINE reports	Cambon, Fouquet, Pelleter, Cathalot	On Owncloud	No Available for LIFEDEEPER consortia
WP 2 and 4	Meiofauna morphology/density data (SO268 & IP21) + Nematoda identifications - collaboration with	SO268: Lefaible et al., 2022 (reviewed) IP21: Lefaible et al., in preparation	https://doi.pangaea.de/10.1594/PANGAEA.942174 https://doi.pangaea.de/10.1594/PANGAEA.942177	No but will be upon publication Nematoda identifications will be added as

	IFREMER (SO268 & IP21			supplementary data
WP4	Demographic and Reproduction data vent shrimp	Hernandez-Avila et al. 2021	https://doi.org/10.17882/84112	Yes
WP2	Biodiversity, isotope & environmental data vents	Sarrazin et al. 2022	https://doi.org/10.17882/88555	Yes
WP2	Stable isotope ratios & elemental contents deep sea	Michel et al. 2020, 2021	https://doi.org/10.17882/76595 https://doi.org/10.17882/62304	Yes
WP4	Population structure & reproductive biology of vent gastropods	Marticorena et al. 2020	https://doi.org/10.17882/71838	Yes
WP2	Temperature data vent habitats	Matabos et al.	https://doi.org/10.17882/80042 https://doi.org/10.17882/80039 https://doi.org/10.17882/80040 https://doi.org/10.17882/52150 https://doi.org/10.17882/52144	Yes
WP2, WP4	Diversity and population structure vents	Marticorena et al. 2020	https://doi.org/10.17882/76246	Yes
WP2	Snake Pit vent assemblages	Cambon et Sarrazin 2020	https://doi.org/10.17882/74349	**Yes**
WP2	Environmental data vent	Sarrazin & Husson 2017	https://doi.org/10.17882/50223 https://doi.org/10.17882/50218	**Yes**
WP2	Taxonomic inventory vent	Sarrazin & Husson 2017	https://doi.org/10.17882/50221	**Yes**
WP2	Image annotations & environmental data vent	Van Audenhaege	https://doi.org/10.17882/84672	**Yes**
WP2	3D reconstructions	Matabos & Arnaubec, 2015	https://doi.org/10.17882/79218	**Yes**
WP2, WP3	Mineralogy, Geochemistry, Isotopic ratios of sediments, hydrothermal deposits, deep sea particles and fluids	Djedjroh, 2022	https://www.theses.fr/s218582	**in progress**

Annex 1

Consent form for interviewee

Insert your institute is the data controller in connection with this research.
We ask for your consent to process your personal data. The purpose of the processing is:
Insert a very brief description of the research activity and the purpose of the data collection

By signing this declaration, you consent to let us: Delete the irrelevant options.

- ❑ Process these personal data: List the data you want to process
- ❑ Publish our results and the following personal data: List the personal data you want to publish The data is published in Enter a short description of the way the data is published.
- ❑ Reuse the data in our research of Enter the research activity/purpose from above in a period of X year(s) after your consent.
- ❑ Disclose the data to our partners in the research of Enter research activity. The research partners share the same purpose and are subject to the same rules of confidentiality, as any insert your partner institute employee. If we disclose your data, you will receive a notice.

The data will be stored securely, and we will use the data solely for the above purpose. You always have the right to revoke your consent. If you wish to revoke your consent later on, you can send an email to: Insert your partner institute's DPO mailbox

Under the General Data Protection Regulation, you are entitled to receive a copy of all your personal information that is held for LIFEDEEPER's research activities by sending a request to Insert project contact's email address.
Best regards,
Project contact
Title or role
Email address

Your signature

☐ (Tick)
I hereby consent to insert partner institute processing my data in accordance with the above purpose and information.
Date:

Name:

Signature