



Series of Technical Reports edited by the CNR-ISMAR DiSSCo Working Group

The Aquatic Natural Collections of the CNR-ISMAR

Guidelines for biological metadata management- v.1

Simona Armeli Minicante, Edoardo Di Russo, Roberta D’Onofrio, Marco Sigovini, Elisa Camatti, Lucilla Capotondi, Ilaria Conese, Luciana Ferraro, Laura Giordano, Valentina Grande, Irene Guarneri, Francesca Maggiore

Overview

The CNR-ISMAR DiSSCo WG developed this manual as part of Activity 6.6 “*National network of the aquatic science collection*” in the Italian Integrated Environmental Research Infrastructures System - ITINERIS Project. This guideline suggests a procedure for managing and harmonising the metadata associated with the physical specimens of the aquatic natural collections preserved at the Institute of Marine Sciences (CNR-ISMAR).

Natural collections are archives containing a large amount of data from various research fields, such as zoology, botany, ecology, biogeography, physiology, geology, paleontology, etc. The metadata harmonisation of physical samples, following international reference standards, and creating a digital sample will make them usable for national and international initiatives, also contributing to the increase of the DiSSCo (Distributed System of Scientific Collections, www.dissco.eu) Italian community in the natural science collections management.

Index

1. Definitions and acronyms	2
2. Aquatic Natural Collections of CNR-ISMAR	3
3. References, ontologies and closed vocabularies	5
4. Metadata and descriptions	6
5. Cataloguing	8
5.1 Image capture and file naming.....	19
5.2 General rules on metadata catalog	20
6. Dataset publication procedures: GRSciColl and GBIF.....	20
7. References.....	24

1. Definitions and acronyms

Biodiversity Information Standards (TDWG): historically known as the Taxonomic Databases Working Group, today's Biodiversity Information Standards (TDWG) is a not-for-profit, scientific and educational association formed to establish international collaboration among the creators, managers and users of biodiversity information and to promote the wider and more effective dissemination and sharing of knowledge about the world's heritage of biological organisms. To achieve its goals, TDWG: develops, ratifies and promotes standards and guidelines for the recording and exchange of data about organisms; acts as a forum for discussing all aspects of biodiversity information management through meetings, online discussions, and publications.

Biological Collection Access Service (BioCAsE): is an international network linking biological collections data from natural history museums, botanical/zoological gardens and research institutions. BioCAsE relies on the Access to Biological Collections Data (ABCD) data exchange standard, which TDWG also administers.

Darwin Core (DwC): is a standard maintained by the Darwin Core maintenance group. It includes a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biological diversity by providing identifiers, labels, and definitions. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

Digital specimen: A 'digital specimen' acts as a surrogate in digital environment for a specific physical specimen, identifying its actual location and authoritatively providing information on its collection event (who, when, where) and taxonomy (what), as well as links to high-resolution images. A digital specimen may expose supplementary information about related literature, traits, tissue samples and DNA sequences, chemical analyses, environmental information, and much more, stored elsewhere than in the NSC itself.

The Distributed System of Scientific Collections (DiSSCo): is a new world-class Research Infrastructure (RI) for Natural Science Collections. The DiSSCo RI aims to create a new business model for one European collection that digitally unifies all European natural science assets, sharing common access, curation, policies and practices across countries while ensuring that all the data complies with the FAIR (Findable, Accessible, Interoperable and Reusable data) principles.

Dublin Core (DC): also known as the Dublin Core™ Metadata Element Set, is a set of fifteen "core" elements (properties) for describing resources.

The Dublin Core™ Metadata Initiative (DCMI): is an organization supporting innovation in metadata design and best practices across the metadata ecology. DCMI works openly, and it is supported by a paid-membership model.

Ecological Metadata Language (EML): is a metadata standard that records information about ecological datasets in a series of modular and extensible XML document types. All of the descriptions of datasets in GBIF.org rely on 'metadata'—that is, the information about data—using the open-source EML standard, which is administered and maintained by The Knowledge Network for Biocomplexity. Each Darwin Core Archive includes as one of its components an EML file (written in XML format).

The Global Biodiversity Information Facility (GBIF): is an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open

access to data about all types of life on Earth. The network draws these diverse data sources together through the use of data standards, including Darwin Core.

Index Herbariorum (IH): is a guide to the world's herbaria. Each institution is assigned a permanent unique identifier in the form of a one to eight letter code. The IH of the institutional CNR-ISMAR Herbarium is "ISMAR".

Italian Integrated Environmental Research Infrastructures System (ITINERIS): is a project funded by EU - Next Generation EU PNRR- Mission 4 "Education and Research" - Component 2: "From research to business" - Investment 3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures". The project is coordinated by the CNR with the main aim to build the network of the Italian environmental Research Infrastructures, connect it to the user community and establish access, through the ITINERIS HUB, to the vast array of knowledge, data, analytical tools, and service produced by the participating IRs.

Metadata: "data about data". This means that metadata are descriptions or information about an object, an entity, a variable or a datapoint, and the primary purpose of metadata is to describe or annotate the data it accompanies in a machine-readable way.

Minimum Information about a Digital Specimen (MIDS): the minimum information elements expected to be present when publishing digitised specimen information. This information is categorised in four levels of digitisation (0-3) according to its type and complexity.

Open nomenclature (ON) qualifiers: the identification's uncertainty or provisional status. The approach is based on an ensemble of terms and their abbreviations (called 'signs'), which can be combined with taxon names. Terms and signs are also called 'qualifiers'. ON qualifiers may express different kinds and degrees of uncertainty.

Physical specimen (or specimen): it represents the basic unit of sampling or investigation; it is relative to a unit of time and space; it is identifiable, can be stored separately and catalogued; it can include one or more specimens or portions thereof (fragments, tissues, sections, DNA, etc.). Depending on the preservation method, the physical sample can be adhered to a stable support (slide, sheet, etc.), or stored in a dry container or immersed in a preservation medium.

Research Organization Registry (ROR): is a global, community-led registry of open persistent identifiers for research organizations. The ROR identifier of CNR-ISMAR is <https://ror.org/02hdf6119>.

Voucher specimen (or voucher): a specimen archived in a permanent collection which serves as physical evidence of its occurrence (time and place) and of any identifications and descriptions based on it. In this context, the voucher is an individual monospecific sample and the physical reference for digitalization purposes and selected based on explicit criteria.

2. Aquatic Natural Collections of CNR-ISMAR

The CNR-ISMAR hosts a large heritage of natural collections relating to different fields of study (botany, zoology, paleontology and geology) (Fig.1) and capable of making a significant contribution to the knowledge of marine biodiversity and geodiversity (see the previous technical report [n.27/2024](#)).

In the framework of the RI DiSSCo (<https://www.dissco.eu/>), the activity core is the transition from Physical Specimens to Digital Specimens, to make the data from Aquatic Natural Collections (ANCs)

available for the community, according to the FAIR principles (Wilkinson et al. 2016). The creation of digital twins of NSC objects originating from research activities, together with maintaining the physical sample, will contribute to detecting biodiversity changes and developing appropriate conservation measures (Nanglu et al. 2023). Additionally, it will enhance information sharing, promoting collaboration, communication, and dissemination within and beyond the scientific community.

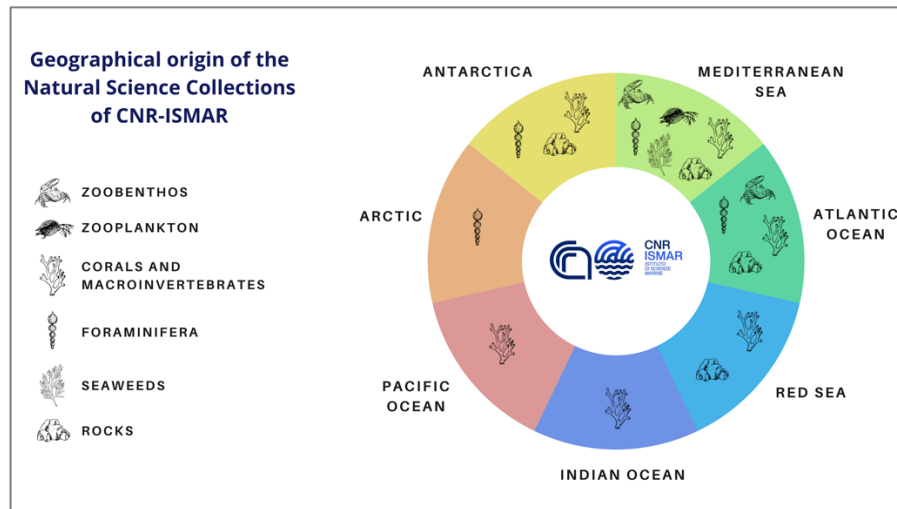


Fig. 1 - Natural collections hosted at the CNR-ISMAR and their geographical origin

In Fig.2 we show the workflow developed by the CNR-ISMAR DiSSCo WG, with the actions needed to organize and make the scientific collections available for research and different stakeholders:

- 1) **curatorship** of physical collections: preservation and maintenance, ensuring that physical specimens are properly stored and preserved to prevent degradation. For example, topping up or replacement of the preservatives, environmental control of the specimens preserved/stored in the slides, control and/or treatment of the herbarium *exsiccata*, checking and/or replacing labels and descriptive sheets, re-organization of the samples in containers, slides, trays, boxes and cabinets, identification of air-conditioned and ventilated rooms to store and/or display the collections;
- 2) specimen **cataloguing** and assignment of the permanent identifier: attributing the catalog number and barcode to each physical object (see paragraph 5);
- 3) **digitisation** and metadating to create digital collections according to international standards: transcription of all available information relating to the physical object into digital format (e.g., excel sheet), image acquisition of specimens and documents to create the digital object of the physical collections;
- 4) **data storage**: maintenance of physical and virtual IT structure to store, organize, and retrieve digital objects;
- 5) **release** of the data and metadata: upload the collection dataset to the Global Biodiversity Information Facility (GBIF) through the Integrated Publishing Toolkit (IPT; see paragraph 6).

In general, after being catalogued, the taxonomic identification phase begins. However, the taxonomic identification is sometimes uncertain or provisional, and the achieved taxonomic resolution may be higher than the species. For this reason, the **identification and/or taxonomic**

revision (*) are actions that do not have a specific time frame but can occur throughout the process.

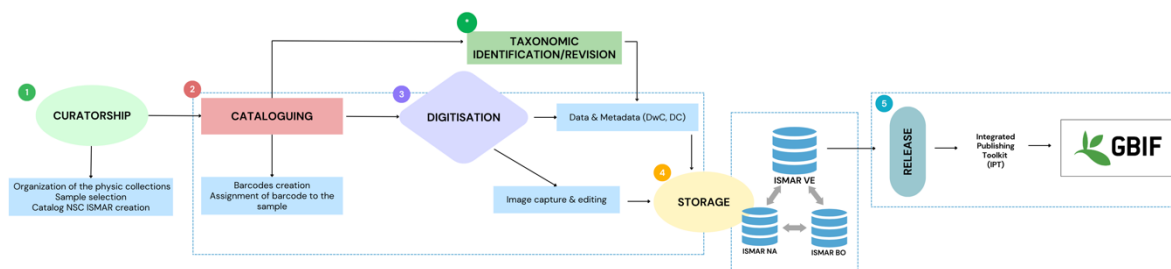


Fig. 2 – Flowchart for managing the aquatic natural collections

This guideline suggests a procedure for managing and harmonising the metadata associated with the physical specimens of the biological natural collections preserved at the Institute of Marine Sciences. According to Dillen et al. (2024), the term ‘data’ is used in the broad sense including both structured, tabular data, such as in databases, and unstructured data, such as notes on specimens and the images of specimens. It also includes the metadata, which makes the data understandable, traceable and reusable.

The data outputs of the digitisation of specimens will be:

- High resolution images of specimen;
- Metadata associated with the imaging of the specimen;
- Unique specimen barcodes.

3. References, ontologies and closed vocabularies

To harmonise the data and metadata vocabulary for each collection, it was decided to refer to closed vocabulary and ontologies, as well as to nomenclature codes, as reported in Tab.1.

Tab.1 – Ontologies, extensions and classification systems used as references

Institution, Specimen, Sampling, Taxonomy	Darwin Core Occurrence Darwin Core Event GBIF Relevé Humboldt Ecological Inventory
Nomenclature	International Code of Nomenclature for algae, fungi, and plants (v.2018) International Code of Zoological Nomenclature (1999)
Spelling and synonyms	WORMS
Open nomenclature	Sigovini et al., 2016 (only for zoological samples)
Biogeography and Habitat	MEOW - Marine Ecoregions of the World EUNIS habitat classification GeoNames
Biological traits	BIOTIC (Biological Traits Information Catalogue) AlgaeTraits
DNA and genomic	DNA derived data GGBN Preparation Extension

International reference standards for metadata	Biodiversity Information Standards (TDWG): Darwin Core Quick Reference Guide (DwC) Minimum Information about a Digital Specimen (MIDS): v0.17
--	--

Note on synonym accepted: the systematic-nomenclatural reference mentioned above is advocated at the time of the final determination of the sample. However, in specific cases, it is possible to use another synonym, with the constraint that the spelling respects the references mentioned above.

4. Metadata and descriptions

Metadata is mainly done using the reference standards of [DarwinCore](#) and [DublinCore](#). To settle the quality of the metadata process, we adopted the *Minimum Information about a Digital Specimen* (MIDS levels; Haston and Hardisty 2020) as quality indicator of the depth of digitisation achieved. As outlined in Table 2, MIDS specifies three levels of ‘minimum information’, together with a pre-level, level zero (0). Level 0 is equivalent to creating a simple catalogue record containing a physical specimen identifier, such as a barcode number. Level 0 often precedes fuller digitisation steps that yield more detailed information. Hence, level 0 is termed a pre-level. Nevertheless, level 0 data is useful minimum information for advertising or knowing about the existence of specimens.

Except for level 0, each MIDS level of minimum information is a superset of the preceding level. The MIDS standard is evolving, and, in this document, we will refer to MIDS Information Elements Jul 2024 version.

Tab. 2 - MIDS levels of minimum information (by Haston et al. 2023 and Jul 2024 version)

Record extent	Purpose	Label	Definition
0 - Bare	A bare or skeletal record making the association between an identifier of a physical specimen and its digital representation, allowing for unambiguous attachment of all other information.	Physical Specimen ID	A unique identity for the specimen within the curating institution. Whatever the institution uses to uniquely identify the item. For example: DOI, stable identifier, catalogue number, barcode, etc.
		Organization	A term to indicate in which institution the specimen is held. This may include an institution code and an institution identifier.
1 - Basic	A basic record of specimen information.	Name	A name given to the object. Any string of characters and/or numbers by which the object is referenced within a collection. This name is not necessarily its name according to an accepted scientific classification, identification, or taxonomic determination.
		Specimen Type	High-level term to delimit and define specimens. For example: preserved specimen, fossil specimen, as opposed to observation.
		Object Type	General term to describe the kind of specimen. In combination with SpecimenType - hierarchical; a more specific classification than described by SpecimenType.

Record extent	Purpose	Label	Definition
2 - Regular	Key information fields that have been agreed over time as essential for most scientific purposes.	License	License under which the specimen data are published
		Modified	UTC date/time of date/time of first creation or subsequent modification, if any
		Qualitative Location	A term to describe the where the specimen was collected. In combination with quantitativeLocation; should capture textual geographic information.
		Quantitative Location	A term to describe where the specimen was collected. A quantitative measure that would include coordinate or shape data, an identifier, or data that can be easily converted into a quantitative measure. In combination with the qualitative term, qualitativeLocality.
		Collecting Agent	A list (concatenated and separated) of names of people, groups, or organizations responsible for recording the original Occurrence.
		Date Collected	The date/time at which a gathering event occurred. For specimen gathering, this is the date/time when the event was recorded.
		Collecting Number	An identifier given to the specimen at the time it was recorded. Often serves as a link between field notes and a specimen record.
		Media	A list (concatenated and separated) of media associated with the specimen.
		Geological Age	Refers to the geological age of a Earth Science specimen (i.e. Fossil, Rock, Mineral or Meteorite) and can be any kind of stratigraphic age, isotopically determined age or structural age.
		Type Status	An indication of the nomenclatural type status of the specimen, where a null value is taken to mean "Assumed not to be a type".
3 - Extended	Other data present or information known about the specimen, including links to third-party sources.	Institution ID	An identifier for the institution having custody of the object(s)
		Collector ID	A list (concatenated and separated) of the identifier(s) for the person, people, groups, or expeditions responsible for collecting the specimen.
		Scientific Name ID	An identifier for the nomenclatural (not taxonomic) details of a scientific name.
		Geographical Locality ID	An identifier for the geographical locality where the specimen was collected.
		Identified By ID	A list (concatenated and separated) of the identifier(s) for the person, people, groups, or organizations responsible for assigning the scientific name to the subject.

To date, 106 terms - grouped into five categories: Basic, Taxonomy, Morphology, Genetic, and Sampling - have been identified for the metadating of the biological natural collections of CNR-ISMAR (Fig.3). The descriptions of the related terms and the reference ontologies are shown in Table 3.

Basic	Taxonomy	Morphological	Genetic	Sampling
occurrenceID	kingdom	organismQuantity	source_mat_ID	eventType
institutionCode	phylum	Morphotype	Preservation medium	eventType name
institutionID	class	Environmental position	preparationType	recordedBy
Institution collection	order	sex	samp_name	eventDate
Catalog code	superfamily	Number of male specimens	single_cell_lysiss_appr	eventTime
collectionCode	family	Number of female specimens	single_cell_lysiss_prot	startDayOfYear
Serie (ID)	subfamily	Reproduction structure	preparationDate	endDayOfYear
datasetID	genus	lifeStage	GenBank Accession Number	Continent
catalogNumber	specificEpithet	measurementType	BOLD Accession Number	Province
basisOfRecord	infraspecificEpithet	measurementValue	associatedSequences	country
type	scientificName	measurementUnit	target_gene	countryCode
modified	taxonRank	Notes	target_subfragment	Ecregion
license	scientificNameAuthorship		pcr_primer_name_forward	stateProvince
accessRights	taxonID		pcr_primer_name_reverse	Region
creator	typeStatus			higherGeography
publisher	identifiedBy			locality
associatedReferences	identifiedByID			verbatimSiteName
bibliographicCitation	Method of identification			Replicate Number
vernacularName	dateIdentified			decimalLatitude
subject (or Keywords)	identificationRemarks			decimalLongitude
otherCatalogNumbers				geodeticDatum
verbatimLabel				verticalDatum
Object Type				Geolocalized geographic object
preparations				coordinateUncertaintyInMeters
verbatimIdentification				verbatimElevation
scientificNameAuthorship				verbatimDepth
ON or other significant qualifier(s)				habitat
Date of first identification				Habitat Classification System
Original Determinavit				eventRemarks
				samplingProtocol

Fig. 3 - Metadata categories and related terms included in the aquatic collections catalogue of the CNR-ISMAR

5. Cataloguing

The **collection code** is the name, acronym, code, or initialism identifying the collection or data set from which the record was derived. As rule established by the working group, **this code must not exceed 6 alphanumeric characters**:

Example: NAdFC

According to Greeff et al. (2021), **catalogue numbers** (DwC:catalogNumber) **should be unique within the respective collection or institution**. To avoid ambiguities with identical numbers used in other collections or institutions, catalogue numbers are often combined with institution codes and collection codes to remain unique on a global scale.

Whitin the ANCs Catalog of CNR-ISMAR, the catalogue number **is composed of the Collection Code followed by a serial number**. As a rule, established by the working group, the serial number does not exceed 9 alphanumeric characters:

Example: NAdFC18E040101

Tab. 3 – Metadata selected to describe the specimens of the Aquatic Natural Collections (in bold: DwC metadata required (**)) and recommended (*) by GBIF)

	Label	Definition	Reference standard / vocabulary	Example
1	occurrenceID	An identifier for the dwc:Occurrence (as opposed to a particular digital record of the dwc:Occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the dwc:occurrenceID globally unique.	DwC	http://arctos.database.museum/guid/MSB:Mamm:233627
2	institutionCode	The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record.	DwC	"Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine", "CNR-ISMAR-VE" "CNR-ISMAR-BO" "CNR-ISMAR-NA"
3	institutionID	An identifier for the institution having custody of the object(s) or information referred to in the record. For physical specimens, the recommended best practice is to use a globally unique and resolvable identifier from a collections registry such as the Research Organization Registry (ROR) or the Global Registry of Scientific Collections (https://www.gbif.org/grscicoll).	DwC	See the paragraph "General rules"
4	Institution collection	Acronym referring to the type of collection (zoology, botany, geology, paleontology).		"BOT", "ZOO", "PAL", "GEO"
5	Catalog code	Acronym referring to the type of collection present at the institute's headquarters. Composed by Institute acronym (ISMAR) + Institute site (VE) + initial of collection field (B = Botany; P = Paleontological; Z = Zoological)		"ISMARVE-Z", "ISMARNA-P", "ISMARBO-B"
6	collectionCode	The name, acronym, code, or initialism identifying the collection or data set from which the record was derived. Note: we have defined a code no longer than 6 alphanumeric characters.	DwC	"DC", "MS", "LicSic"
7	Serie (ID)	The identification code of a series within the collection. It may not always be present.		"Miscellanea", "AS", "XiLag", "LicSic17"
8	datasetID	An identifier for the set of data. May be a global unique identifier or an identifier specific to a collection or institution.	DwC	"b15d4952-7d20-46f1-8a3e-556a512b04c5"
9	catalogNumber	An identifier (preferably unique) for the record within the data set or collection. NOTE: the code is composed from the Collection Code + the Serial number (consisting of a maximum of 9 alphanumeric characters).	DwC	"DC00213"
10	basisOfRecord **	The specific nature of the data record. Recommended best practice is to use this controlled vocabulary: http://rs.gbif.org/vocabulary/dwc/basis_of_record.xml	DwC	"PreservedSpecimen", "FossilSpecimen", "LivingSpecimen", "HumanObservation", "MachineObservation"
11	type	The nature or genre of the resource. Must be populated with a value from the DCMI type vocabulary	DwC	"StillImage", "MovingImage", "Sound", "PhysicalObject", "Event", "Text"

		(https://www.dublincore.org/specifications/dublin-core/dcmi-type-vocabulary/2010-10-11/)		
12	modified	The most recent date-time on which the resource was changed. http://purl.org/dc/terms/modified	DwC	"AAAA-MM-GG"
13	License	A legal document giving official permission to do something with the resource.	DwC	CC-BY 4.0
14	accessRights	Information about who can access the resource or an indication of its security status. http://purl.org/dc/terms/accessRights	DwC	"License CC-BY 4.0 with obligation to cite the Institute of Marine Sciences, National Research Council (CNR-ISMAR)"
15	creator	An entity responsible for making the resource. According to GBIF: "One or more individuals, groups, or institutions responsible for the creation of the dataset". https://ipt.gbif.org/manual/en/ipt/latest/citation	DwC	"Mario Rossi Donatella Bianchi"
16	publisher	An entity responsible for making the resource available.	DwC	"CNR-ISMAR, Atlas of Aquatic Natural Collections"
17	associatedReferences	A list (concatenated and separated) of identifiers (publication, bibliographic reference, global unique identifier, URI) of literature associated with the dwc:Occurrence. http://rs.tdwg.org/dwc/terms/associatedReferences	DwC	"Steven R. Hoofer and Ronald A. Van Den Bussche. 2001. Phylogenetic Relationships of Plecotine Bats and Allies Based on Mitochondrial Ribosomal Sequences. Journal of Mammalogy 82(1):131-137. Walker, Faith M., Jeffrey T. Foster, Kevin P. Drees, Carol L. Chambers. 2014. Spotted bat (Euderma maculatum) microsatellite discovery using illumina sequencing. Conservation Genetics Resources"
18	bibliographicCitation	A bibliographic reference for the resource. The intended usage of this term in Darwin Core is to provide the preferred way to cite the resource itself - "how to cite this record". Note that the intended usage of dcterms:references in Darwin Core, by contrast, is to point to the definitive source representation of the resource - "where to find the as-close-to-original reference, if one is available.	DwC	"Museum of Vertebrate Zoology, UC Berkeley. MVZ Mammal Collection (Arctos). Record ID: http://arctos.database.museum/guid/MVZ:Mamm:165861?seid=101356 . Source: http://ipt.vertnet.org:8080/ipt/resource.do?r=mvz_mammal ", " https://www.gbif.org/species/2439608 Source: GBIF Taxonomic Backbone", "Rand, K.M., Logerwell, E.A. The first demersal trawl survey of benthic fish and invertebrates in the Beaufort Sea since the late 1970s. Polar Biol 34, 475–488 (2011). https://doi.org/10.1007/s00300-010-0900-2 "
19	vernacularName	A common or vernacular name.	DwC	"Vongola filippina", "lattuga di mare", "Ostrica"
20	subject (or Keywords)	The topic of the resource. A list (concatenated and separated) of keywords useful for the user to search for samples in the repository.	DC	"Ulva californica Chlorophyta Venice Lagoon alien species"
21	otherCatalogNumbers	A list (concatenated and separated) of previous or alternate fully qualified catalog numbers or other human-used identifiers for the same dwc:Occurrence, whether in the current or any other data set or collection.	DwC	"NPS YELLO6778 MBG 33424"

22	verbatimLabel (name)	The content of this term should include no embellishments, prefixes, headers or other additions made to the text. Abbreviations must not be expanded and supposed misspellings must not be corrected. Lines or breakpoints between blocks of text that could be verified by seeing the original labels or images of them may be used. Examples of material entities include preserved specimens, fossil specimens, and material samples. Best practice is to use UTF-8 for all characters. Best practice is to add comment "verbatimLabel derived from human transcription" in dwc:occurrenceRemarks. NOTE: Pre-1980 labels must be kept as is. Subsequent labels may be subject to replacement for the purposes of their final cataloging. If multiple labels, they are all inserted in chronological order (for example "common oyster Ostrea edulis"). The field is descriptive, it has no length constraints.	DwC	"Ostrea edulis", "Ostrea", "Mollusca", "Fungo"
23	Object Type	General term to describe the kind of specimen. In combination with SpecimenType - hierarchical; a more specific classification than described by SpecimenType.		"microscope slide", "pinned insect", "herbarium sheet"
24	preparations	Medium or support in which the specimen is preserved. A list (concatenated and separated) of preparations and preservation methods for a dwc:MaterialEntity.	DwC	"denatured alcohol"; "ethyl alcohol 90%"; "ethyl alcohol 70%"; "frozen"; "dry"; "pressed and dried"; "formalin"; "formalin 4%"; "rose bengal 2gr/l"
25	verbatimIdentification	A string representing the taxonomic identification as it appeared in the original record. This term is meant to allow the capture of an unaltered original identification/determination, including identification qualifiers, hybrid formulas, uncertainties, etc. This term is meant to be used in addition to dwc:scientificName (and dwc:identificationQualifier etc.), not instead of it.	DwC	"Peromyscus sp.", "Pachyporidae?", "Potentilla × pantotricha Soják"
26	scientificNameAuthorship	The authorship information for the dwc:scientificName formatted according to the conventions of the applicable dwc:nomenclaturalCode.	DwC	BOT: "Linnaeus", "(L.) W. D. J. Koch" ZOO: "Linnaeus, 1758"
27	Open Nomenclature or other significant qualifier(s)	Open Nomenclature is a semantic tool that allows scientists to communicate the uncertainty of a taxonomic identification by means of qualifiers combined to taxon names. ON qualifiers can be used as a provisional notation or may be included definitively in taxonomic lists and SA matrices (Sigovini et al. 2016).	Open Nomenclature	"sensu de Azevedo Ferreira & Tavares (2020)"
28	Date of first identification	The date on which the subject was determined. Recommended best practice is to use a date that conforms to ISO 8601-1:2019.		"AAAA-MM-GG", "AAAA", "AAAA-MM"
29	Original Determinavit	A list (concatenated and separated) of names of people, groups, or organizations who assigned the dwc:Taxon to the subject.		"Aristocle Vatova Michelangelo Minio"

30	kingdom *	The full scientific name of the kingdom in which the dwc:Taxon is classified.	DwC	"Animalia", "Archaea", "Bacteria", "Chromista", "Fungi", "Plantae", "Protozoa", "Viruses"
31	phylum	The full scientific name of the phylum or division in which the dwc:Taxon is classified.	DwC	"Mollusca"
32	class	The full scientific name of the class in which the dwc:Taxon is classified.	DwC	"Bivalvia"
33	order	The full scientific name of the order in which the dwc:Taxon is classified.	DwC	"Ostreida"
34	superfamily	The full scientific name of the superfamily in which the dwc:Taxon is classified.	DwC	"Pinnoidea"
35	family	The full scientific name of the family in which the dwc:Taxon is classified.	DwC	"Pinnidae"
36	subfamily	The full scientific name of the subfamily in which the dwc:Taxon is classified.	DwC	
37	genus	The full scientific name of the genus in which the dwc:Taxon is classified.	DwC	"Pinna"
38	specificEpithet	The name of the first or species epithet of the dwc:scientificName.	DwC	"nobilis"
39	infraspecificEpithet	The name with the lowest or terminal infraspecific epithet of the scientificName excluding any rank designation.	DwC	"Homo sapiens sapiens", "Ulva flexuosa subsp. flexuosa"
40	scientificName **	The full scientific name, with authorship and date information if known. When forming part of a dwc:Identification, this should be the name in lowest level taxonomic rank that can be determined. This term should not contain identification qualifications, which should instead be supplied in the dwc:identificationQualifier term.	DwC	"Pinna nobilis"
41	taxonRank	The taxonomic rank of the most specific name in the dwc:scientificName.	DwC	subspecies, varietas, forma, species, genus
42	scientificNameAuthorship	The authorship information for the dwc:scientificName formatted according to the conventions of the applicable dwc:nomenclaturalCode.	DwC	BOT: "Linnaeus", "(L.) W. D. J. Koch" ZOO: "Linnaeus, 1758"
43	taxonID	An identifier for the set of dwc:Taxon information. May be a global unique identifier or an identifier specific to the data set. NOTE: we refer to AphiaID.	DwC	"148592"
44	typeStatus	A list (concatenated and separated) of nomenclatural types (type status, typified scientific name, publication) applied to the subject. Recommended best practice is to separate the values in a list with space vertical bar space (). This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.	DwC	"holotype of Picea abies"

45	identifiedBy	A list (concatenated and separated) of names of people, groups, or organizations who assigned the dwc:Taxon to the subject. Recommended best practice is to separate the values in a list with space vertical bar space (). This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.	DwC	"Theodore Pappenfuss Robert Macey"
46	identifiedByID	A list (concatenated and separated) of the globally unique identifier for the person, people, groups, or organizations responsible for assigning the dwc:Taxon to the subject.	DwC	https://orcid.org/0000-0002-1825-0097 https://orcid.org/0000-0002-1825-0098
47	Method of identification	Method of identification applied.		"Classical taxonomy", "molecular taxonomy", "AI", "other"
48	dateIdentified	The date on which the subject was determined as representing the dwc:Taxon. Recommended best practice is to use a date that conforms to ISO 8601-1:2019.	DwC	"AAAA-MM-GG", "AAAA", "AAAA-MM"
49	identificationRemarks	Comments or notes about the dwc:Identification.	DwC	
50	organismQuantity	A number or enumeration value for the quantity of dwc:Organisms.	DwC	"27 individuals", "12.5 % biomass", "many individuals"
51	Morphotype	The morphotype of the observed specimen. NOTE: Since morphotypes can differ between organism groups, provide the reference of the field Notes.	BIOTIC AlgaeTraits	"Articulate", "Crustose"
52	Environmental position	Position relative to substratum or fluid medium (air/water)	AlgaeTraits	"Endolithic", "Endophytic", "Endozoic", "Epilithic", "Epiphytic", "Epizoic", "Unattached", "Episamnic"
53	sex	The sex of the biological individual(s) represented in the dwc:Occurrence.	DwC	"Male", "Female", "Male and female", "Hermaphrodite", "Undetermined"
54	Number of male specimens	Number of male specimens		"3", "56"
55	Number of female specimens	Number of female specimens		"3", "56"
56	Reproduction structure	Presence of reproduction structure		"Yes", "No"
57	lifeStage	The age class or life stage of the dwc:Organism(s) at the time the dwc:Occurrence was recorded.	DwC	"Juvenile", "Adult", "Larva", "Unknown", "Sporophyte", "Gametophyte", "Spore", "Tetrasporophyte"
58	measurementType	The nature of the measurement, fact, characteristic, or assertion.	DwC	"Length"
59	measurementValue	The value of the measurement, fact, characteristic, or assertion.	DwC	"15"
60	measurementUnit	The units associated with the dwc:measurementValue.	DwC	"cm"
61	Notes	Comments/notes about the morphology, and/or specific metadata		
62	source_mat_ID	A unique identifier assigned to a material sample used for extracting nucleic acids, and subsequent sequencing. The identifier can refer either to the original material collected or to any derived sub-samples	DwC	"MPI012"

63	Preservation medium	Medium in which the tissue fragment from which the DNA is extracted is preserved		"Silica gel", "herbarium sheet", "70% alcohol", "slide", "dry sediment"
64	preparationType	Description of preparation type (specimens, tissues, DNA).	DwC	for DNA: gDNA; for tissues/specimens: leaf, muscle, leg, blood
65	samp_name	Sample Name is a name that you choose for the sample. It can have any format, but we suggest that you make it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique. The code attributed to the Tissue during the DNA extraction procedure	DwC	"U1", "03"
66	single_cell_lysis_appr	Method used to free DNA from interior of the cell(s) or particle(s). Recommended best practice is to use this controlled vocabulary: https://rs.gbif.org/vocabulary/dna/single_cell_lysis_appr.xml	DwC	"chemical", "enzymatic", "physical", "combination"
67	single_cell_lysis_prot	Name of the kit or standard protocol used for cell(s) or particle(s) lysis	DwC	"DNeasy Qiagen", "Saunders and McDevit 2012", ...
68	preparationDate	The date of preparation/extraction	DwC	"AAAA-MM-GG"
69	GenBank Accession Number	The unique code issued by GenBank and attributed to the deposited sequence		"MF172090"
70	BOLD Accession Number	The unique code issued by BOLDSystem and attributed to the deposited sequence		"ITGRE008-11"
71	associatedSequences	A list (concatenated and separated) of identifiers (publication, global unique identifier, URI) of genetic sequence information associated with the dwc:MaterialEntity.	DwC	"http://www.ncbi.nlm.nih.gov/nuccore/GU328060 http://www.ncbi.nlm.nih.gov/nuccore/AF326093"
72	target_gene	Targeted gene or locus name for marker gene study	DwC	16S rRNA, 18S rRNA, COI
73	target_subfragment	Name of subfragment of a gene or locus. Important to e.g. identify special regions on marker genes like V6 on 16S rRNA	DwC	V6, V9, ITS
74	pcr_primer_name_forward	Name of the forward PCR primer that were used to amplify the sequence of the targeted gene, locus or subfragment. If multiple multiple forward or reverse primers are present in a single PCR reaction, there should be a full row for each of these linked to the same DWC Occurrence.	DwC	jgLCO1490
75	pcr_primer_name_reverse	Name of the reverse PCR primer that were used to amplify the sequence of the targeted gene, locus or subfragment. If multiple multiple forward or reverse primers are present in a single PCR reaction, there should be a full row for each of these linked to the same DWC Occurrence.	DwC	jgHCO2198

76	eventType	The nature of the dwc:Event. Recommended best practice is to use a controlled vocabulary. Regardless of the dwc:eventType, the interval of the dwc:Event can be captured in dwc:eventDate. This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.	DwC	Sample, Observation, Site Visit, Biotic Interaction, Bioblitz, Expedition, Survey, Project
77	eventType name	Name of the project or survey in which the sample was collected.	DwC	"RITMARE"
78	recordedBy (or collectingAgent, or legit) 2L	A list (concatenated and separated) of names of people, groups, or organizations responsible for recording the original Occurrence.	DwC	"Aristocle Vatova"
79	eventDate **	The date-time or interval during which a dwc:Event occurred. For occurrences, this is the date-time when the dwc:Event was recorded. Not suitable for a time in a geological context. Recommended best practice is to use a date that conforms to ISO 8601-1:2019.	DwC	"AAAA-MM-GG"
80	eventTime	The time or interval during which a dwc:Event occurred. Recommended best practice is to use a time of day that conforms to ISO 8601-1:2019.	DwC	"13:00:00Z/15:30:00Z"
81	startDayOfYear	The earliest integer day of the year on which the dwc:Event occurred (1 for January 1, 365 for December 31, except in a leap year, in which case it is 366).	DwC	1 (1 January)
82	endDayOfYear	The latest integer day of the year on which the dwc:Event occurred (1 for January 1, 365 for December 31, except in a leap year, in which case it is 366).	DwC	32 (1 February)
83	Continent	The name of the continent in which the dcterms:Location occurs.	DwC	"Africa", "Antarctica", "Asia", "Europe", "North America", "Oceania", "South America"
84	Province	Large areas defined by the presence of distinct biotas that have at least some cohesions over evolutionary time frames. Provinces will hold some level of endemism, principally at the level of species. Although historical isolation will play a role, many of these distinct biotas have arisen as a result of distinctive abiotic features that circumscribe their boundaries. These may include geomorphological features (isolated island and shelf systems, semienclosed seas); hydrographic features (currents, upwellings, ice dynamics); or geochemical influences (broadest-scale elements of nutrient supply and salinity).	MEOW	Arctic, Northern European Seas, Lusitanian, Mediterranean Sea, Cold Temperate Northwest Atlantic, Warm Temperate Northwest Atlantic, Black Sea
85	country	The name of the country or major administrative unit in which the dcterms:Location occurs.	DwC	"Italy", "Chile", "Australia"
86	countryCode **	A unique (preferably globally-unique) identifier for the taxon represented in the row. Recommended best practice is to use ISO 3166-1-alpha-2 country codes: http://rs.gbif.org/vocabulary/iso/3166-1_alpha2.xml	DwC	"AR", "SV"

87	Ecoregion	Areas of relatively homogeneous species composition, clearly distinct from adjacent systems. The species composition is likely to be determined by the predominance of a small number of ecosystems and/or a distinct suite of oceanographic or topographic features. The dominant biogeographic forcing agents defining the eco-regions vary from location to location but may include isolation, upwelling, nutrient inputs, freshwater influx, temperature regimes, ice regimes, exposure, sediments, currents, and bathymetric or coastal complexity.	MEOW	South and West Iceland, Faroe Plateau, Southern Norway, Baltic Sea, Adriatic Sea, Levantine Sea, Northern Gulf of Mexico, Black Sea
88	stateProvince	The name of the next smaller administrative region than country (state, province, canton, department, region, etc.) in which the dcterms:Location occurs.	DwC	"Montana", "Cordoba"
89	Region	The name of the administrative region in which the dcterms:location occurs.	GeoNames	"Liguria", "Tuscany"
90	higherGeography	A list (concatenated and separated) of geographic names less specific than the information captured in the dwc:locality term.	DwC	"Europe Mediterranean Sea Adriatic Sea Venice Lagoon"
91	locality	The specific description of the place. Less specific geographic information can be provided in other geographic terms (dwc:higherGeography, dwc:continent, dwc:country, dwc:stateProvince, dwc:county, dwc:municipality, dwc:waterBody, dwc:island, dwc:islandGroup). This term may contain information modified from the original to correct perceived errors or standardize the description.	DwC	"Punta Vela", "Palude della Rosa"
92	verbatimSiteName	The sampling station code. A list (concatenated and separated) of original site names.	Humboldt Ecological Inventory	"ZN1", "ZN1_03", "PL1_180"
93	Replicate Number	The number of any sampling replicate.		"SAM01.1", "SAM01.2"
94	decimalLatitude **	The geographic latitude (in decimal degrees, using the spatial reference system given in dwc:geodeticDatum) of the geographic center of a dcterms:Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive.	DwC	"45,21349"
95	decimalLongitude **	The geographic longitude (in decimal degrees, using the spatial reference system given in dwc:geodeticDatum) of the geographic center of a dcterms:Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.	DwC	"12,2745"

96	geodeticDatum *	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude are based. Recommended best practice is use the EPSG code as a controlled vocabulary to provide an SRS, if known. Otherwise use a controlled vocabulary for the name or code of the geodetic datum, if known. Otherwise use a controlled vocabulary for the name or code of the ellipsoid, if known. If none of these is known, use the value "unknown".	DwC	"WGS84", "EPSG:4326", "NAD27", "Campo Inchauspe", "European 1950", "Clarke 1866"
97	verticalDatum	The vertical datum used as the reference upon which the values in the elevation terms are based. Recommended best practice is to use a controlled vocabulary. This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.	DwC	"EPSG:1051"
98	Geolocalized geographic object			"Sampling station", "Locality centroid"
99	coordinateUncertaintyInMeters	The horizontal distance (in meters) from the given dwc:decimalLatitude and dwc:decimalLongitude describes the smallest circle containing the whole of the dcterms:Location. Leave the value empty if the uncertainty is unknown, cannot be estimated, or is not applicable (because there are no coordinates). Zero is not a valid value for this term.	DwC	"100"
100	verbatimElevation	The original description of the elevation (altitude, usually above sea level) of the Location.	DwC	"200m"
101	verbatimDepth	The original description of the depth below the local surface.	DwC	"64m"
102	habitat	A category or description of the habitat in which the dwc:Event occurred. This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.	DwC	"X03 - Brackish coastal lagoons", "intertidal", "coralligenous"
103	Habitat Classification System	The reference used for the field dwc:Habitat		"EUNIS 2012 (2019)", "IUCN Red list v.3.1"
104	eventRemarks	Comments or notes about the dwc:Event.	DwC	

105	samplingProtocol	The names of, references to, or descriptions of the methods or protocols used during a dwc:Event. Recommended best practice is describe a dwc:Event with no more than one sampling protocol. In the case of a summary Event with multiple protocols, in which a specific protocol cannot be attributed to specific dwc:Occurrences, the recommended best practice is to separate the values in a list with space vertical bar space (). This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.	DwC	"UV light trap", "Box-corer", "ISPRA, Marine Strategy", "SIBM..."
-----	-------------------------	--	---------------------	---

5.1 Image capture and file naming

For image acquisition, a resolution of no less than 400 dpi is recommended, and in any case between 400 and 600 dpi. Given its information preservation characteristics, the format chosen for file storage is TIFF, .tif extension, uncompressed and in colour (RGB) (see Armeli Minicante et al. 2017).

This image is not usually published but serves as the base from which other image versions are derived (e.g., .jpeg extension). The resolution of these images will vary according to the intended use of the derived images (such as web publishing or printing) (Nieva de la Hidalga et al. 2020).

According to Nieva de la Hidalga et al. (2020) and reference therein, “*images must include a set of visual elements that appear next to the specimen and are intended to help in the identification, processing and quality control.*” These elements include: color checker (A), scale bar (B), labels (C), QRcode or barcode (D) and Institution name (E).

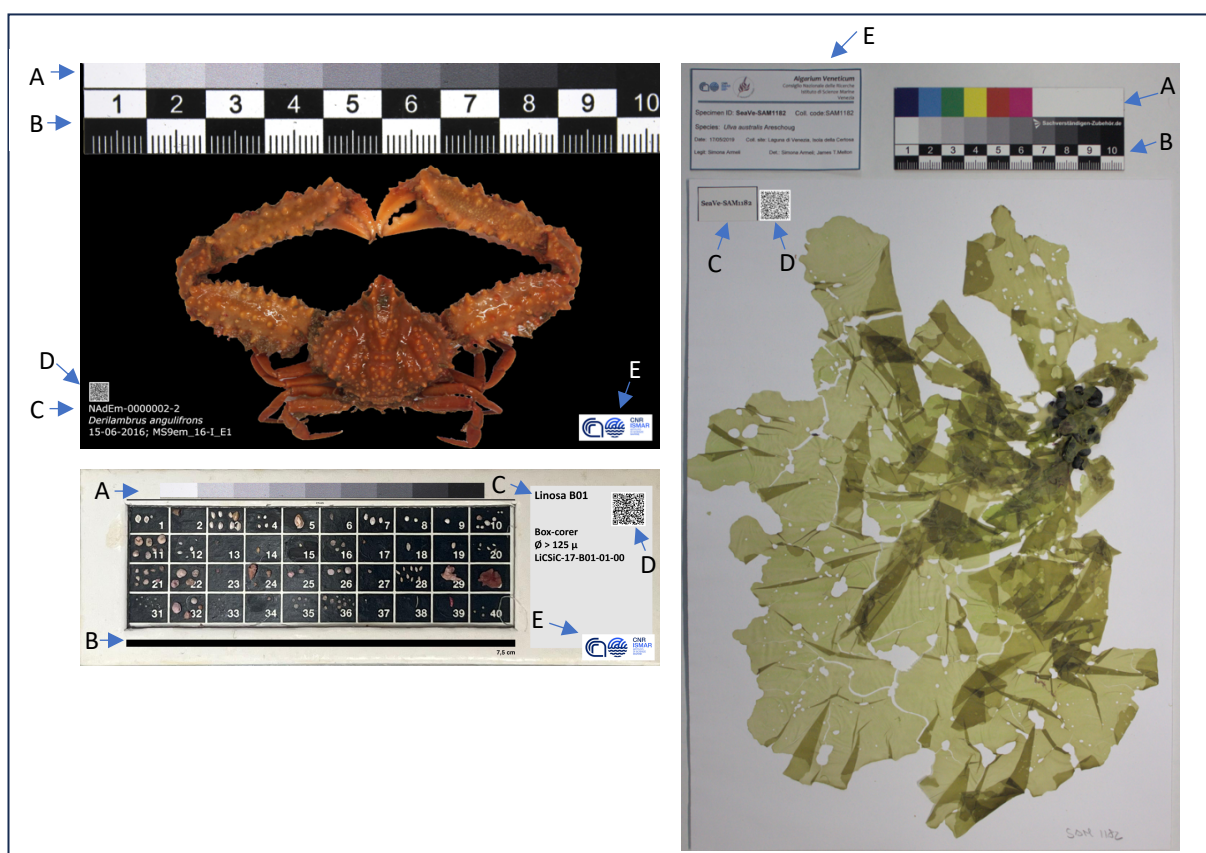


Fig. 4 – Examples of specimen images including the visual elements: color checker (A), scale bar (B), labels (C), QRcode (D) and Institution name (E)

The **color checker** is recommended by the Federal Agencies Guideline for Digitisation (FADGI, <https://www.digitizationguidelines.gov/>) for helping with quality control and post-processing; this can help in verifying the lighting, white balance and color accuracy of the image. **Scale bar** is recommended to enable the calculation of the dimensions of the specimen, whereas, the **institution name** (with or without logo) is required to quickly identify the institution holding the specimen (Phillips et al. 2014). **Labels** are commonly placed next to the specimen. Concerning herbarium samples, a clear capture of labels is important for further processing and documentation of the

specimens (Nieva de la Hidalga et al. 2020). **Barcodes** (or QRcodes) are identifiers used for cataloguing specimens which are also useful for linking them to digital specimens.

All image files from the same specimen must be named with the catalog code:

Example: NAdFC18E040101

Multiple image files of the same sample should be named sequentially and, if possible, reporting the magnification of the lens:

Example: NAdFC18E040101_1; NAdFC18E040101_2; NAdFC18E040101_3(40x); ...;

5.2 General rules on metadata catalog

- Do not write in italics or bold.
- Do not use brackets (except for the Authors of the species according to the nomenclature codes).
- In case of codes (e.g. sampling stations), avoid spaces and use underscores (e.g. ST01_20).
- Enter dates in the format YYYY-MM-DD.
- Do not use the string 'N/A' when a value is unknown.
- scientificName should not contain any remarks. [For example, 'Quinqueloculina juvenile specimens' isn't right; the correct field of 'juvenile' is [lifeStage](#). GBIF interprets 'juvenile specimen' as specific and intraspecific epithets when in the field scientificName.].
- It's a good practice to use the GBIF's species matching tool, to check the match between your names and those in GBIF's taxonomic backbone: <https://www.gbif.org/tools/species-lookup>.
- Make sure to have not unicode no-break space at the end of the text.

6. Dataset publication procedures: GRSciColl and GBIF

To publish the ISMAR collection datasets - in accordance with FAIR principles - we utilized the Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>) and the GBIF Registry of Scientific Collections builds on the former Global Registry of Scientific Collections (GRSciColl, <https://scientific-collections.gbif.org/>).

GRSciColl is a clearinghouse of information about the world's scientific institutions, collections and associated staff members. It currently holds information on more than 8000 collections, more than 8000 institutions and 15,000 staff members. The principal objectives of the GBIF registry are twofold. Firstly, access to information about institutions, their scientific collections and staff members should be facilitated, and secondly, the GBIF registry provides machine-readable identifiers and unique codes for institutions and collections, elements used in the Darwin Core standard data exchange format by the biodiversity informatics community. These codes are also used for citations in journals such as ZooKeys. Other registries such as Index Herbariorum, the NCBI Institution table and Collection table, Biorepositories.org or the Biodiversity Collections Index are included in the GBIF registry as well. The GBIF registry of scientific collections seeks to preserve historical identifiers for institutions and collections and provides resolution services for codes.

GBIF is an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth. Coordinated through its Secretariat in Copenhagen, the GBIF network of participating countries and organizations, working through the participant nodes, provides data-holding institutions around the world with common standards, best practices and open-source tools enabling them to share information about where and when species have been recorded. This knowledge derives from many different kinds of

sources, including everything from museum specimens collected in the 18th and 19th century to DNA barcodes and smartphone photos recorded in recent days and weeks. The network draws these diverse data sources together through the use of data standards, including Darwin Core, which forms the basis for the bulk of GBIF.org's index of hundreds of millions of species occurrence records. Publishers provide open access to their datasets using machine-readable Creative Commons licence designations, allowing scientists, researchers and others to apply the data in nearly five peer-reviewed publications every day, along with other reports, analyses and policy documents.

Currently, CNR-ISMAR is registered as Publisher Institution on GBIF, while 3 of its branches (Venezia, Bologna and Napoli) are registered on GRSciColl as locations where physical collections are hosted (Table 4).

Tab. 4 – GRSciColl identifiers of the ISMAR branches registered

GrSciColl identifiers			
Name	Istituto di Scienze Marine di Venezia	Istituto di Scienze Marine di Napoli	Istituto di Scienze Marine di Bologna
Code	CNR-ISMAR-VE	CNR-ISMAR-NA	CNR-ISMAR-BO
Additional names	Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine	CNR-ISMAR of Naples	CNR-ISMAR of Bologna
UUID	7b70ff3e-6a7a-4fcf-8754-11c77a15fc3f	9a465359-cf9b-4dc1-86b6-4351d075ae97	f066f8a5-6331-48d1-a9eb-01cfe6f7c969

Once the dataset catalog has been compiled by the researcher/curator of the collection as described above, it is reviewed by the GBIF technical point of ISMAR. If the dataset catalog is related to a naturalistic collection, the collection is first registered under the institution of affiliation on the GRSciColl platform by the GRSciColl Editor User. Following the registration, the obtained collection UUID (Universally Unique Identifier) will be inserted within the dataset catalog as the DwC:datasetID. Via access to the IPT (Integrated Publishing Toolkit), the GBIF technical point of ISMAR upload the dataset catalog, mapping the metadata and providing it with information requested by GBIF. Once published, the administrator will validate the dataset by registering it on GBIF and releasing the related DOI. As a result, each dataset published on GBIF will have a DOI and a reference citation (Fig.5); furthermore, containing the collection UUID in its metadata specimen records are automatically associated with the collection registered on GRSciColl and consequently with the relative branche (Fig.6).

The procedure to publish the Aquatic Natural Collections dataset on GBIF and GRSciColl is showed in Fig.7. To date, 7 dataset collection have been registered on GBIF, for a total of 2,012 occurrences (Fig.8 and Tab.5).

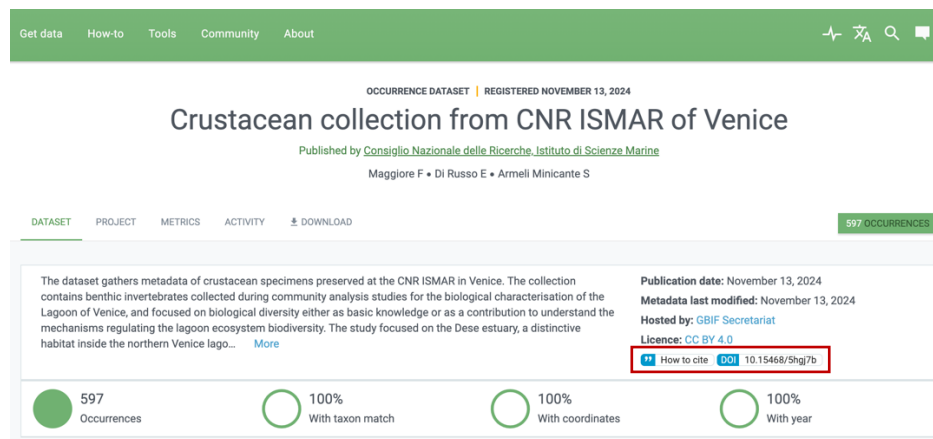


Fig. 5 - Dataset page of a collection registered on the GBIF platform, containing the link to the citation and the DOI of the dataset

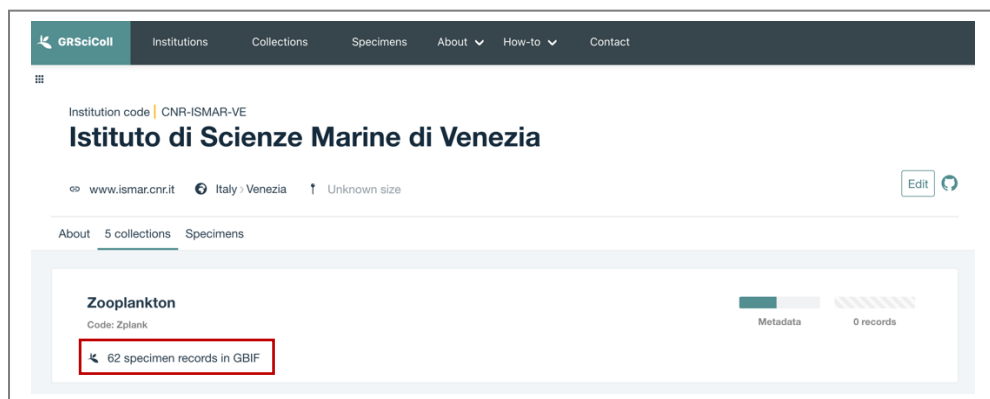


Fig. 6 – Institution page on GRSciColl reporting the specimen records of the collection registered on GBIF

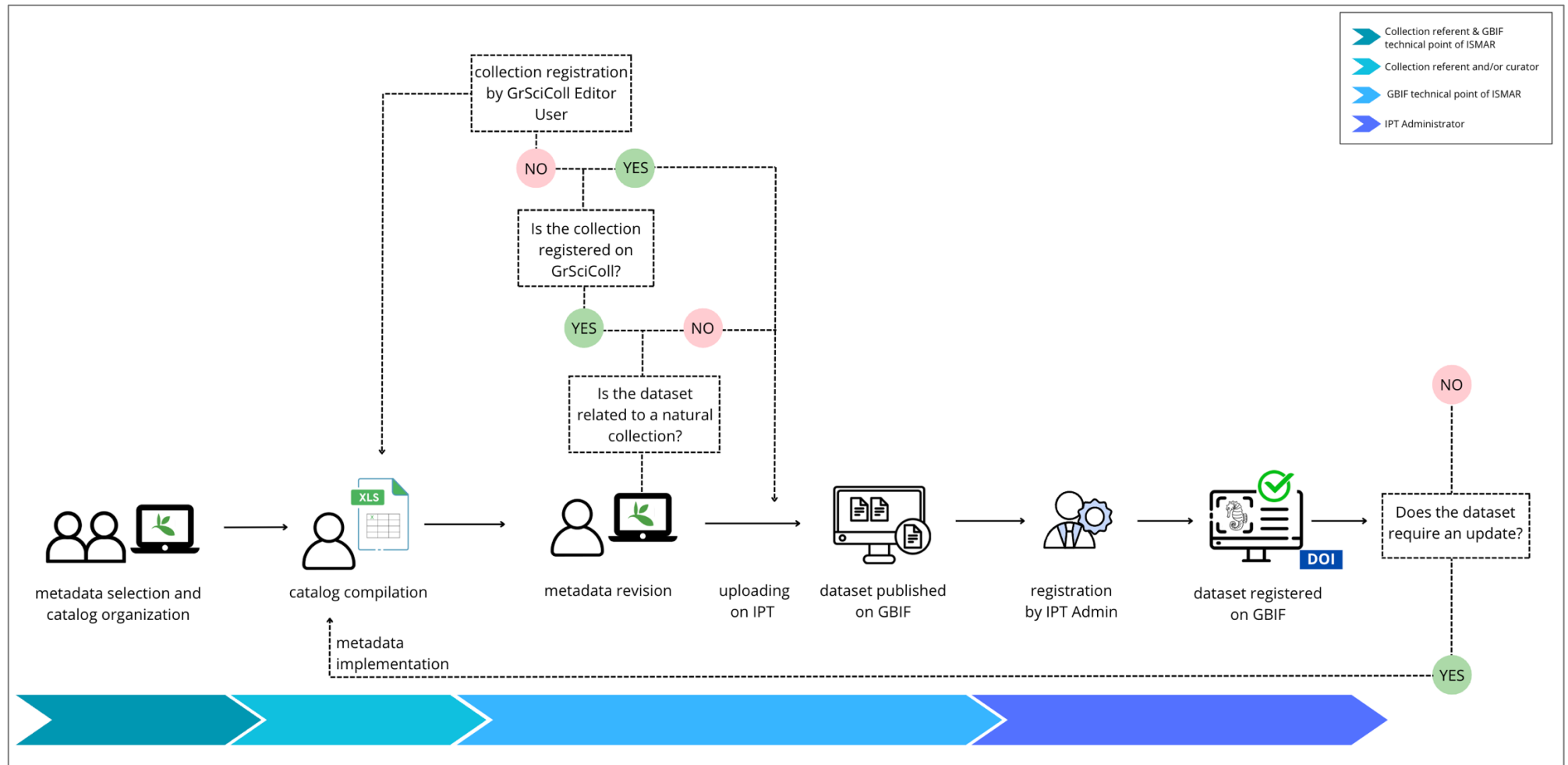


Fig. 7 – Procedure to publish the Aquatic Natural Collections dataset on GBIF and GRSciColl

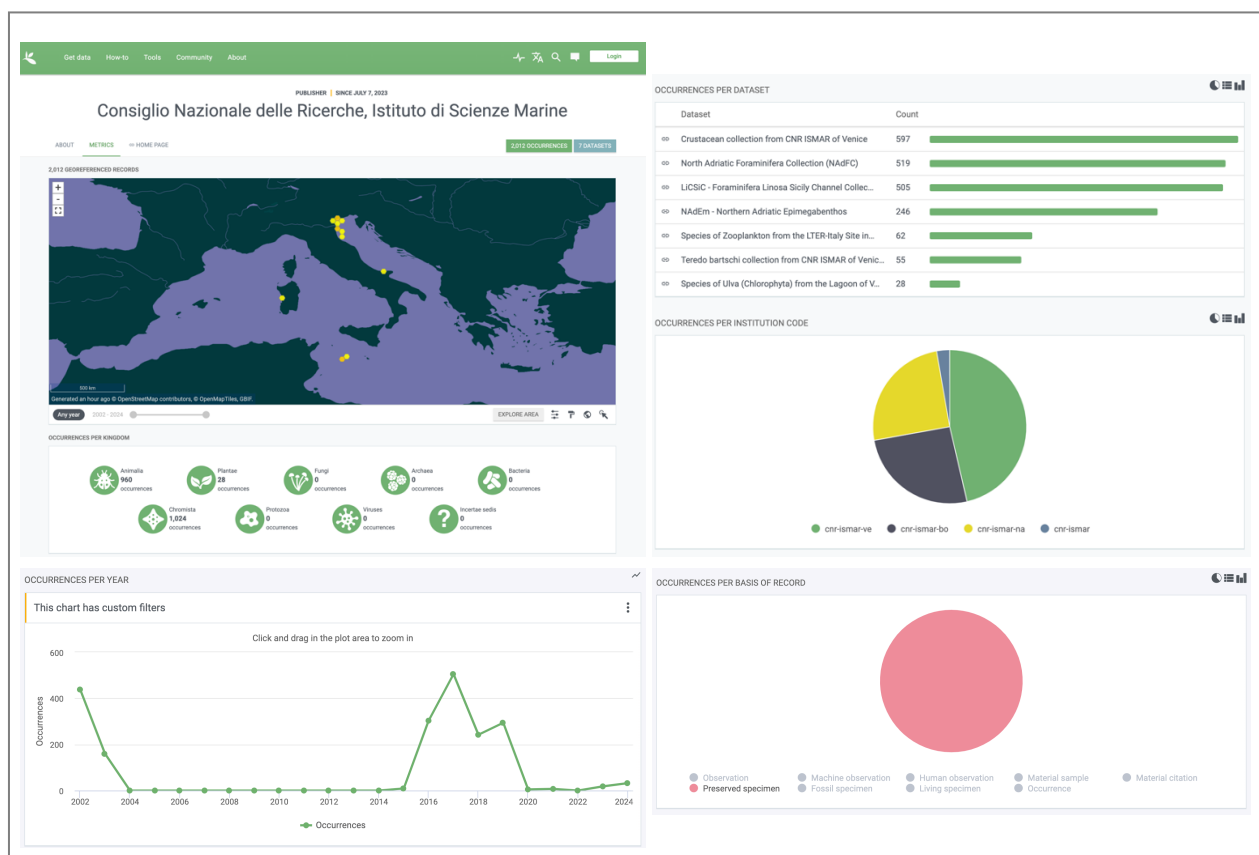


Fig. 8 – ISMAR page metrics on GBIF platform (access: 3 December 2024)

Tab. 5 – Occurrence dataset published on GBIF

Occurrence dataset	DOI
Maggiore F, Di Russo E, Armeli Minicante S (2024). Crustacean collection from CNR ISMAR of Venice.	https://doi.org/10.15468/5hgi7b
Giordano L, Ferraro L (2024). LiCSiC - Foraminifera Linosa Sicily Channel Collection	https://doi.org/10.15468/gv8k7a
Sigovini M, Guarneri I, Tagliapietra D, Sabino A, Di Russo E (2024). NAdEm - Northern Adriatic Epimegabenthos	https://doi.org/10.15468/r8f2c2
D'Onofrio R, Capotondi L (2024). North Adriatic Foraminifera Collection (NAdFC)	https://doi.org/10.15468/2b5bha
Camatti E, Di Russo E, Nurra N (2024). Species of Zooplankton from the LTER-Italy Site in the Lagoon of Venice and Oceanographic Platform CNR (North Adriatic Sea)	https://doi.org/10.15468/8scg2m
Armeli Minicante S, Di Russo E, Lopez-Bautista J (2024). Species of <i>Ulva</i> (Chlorophyta) from the Lagoon of Venice (Italy)	https://doi.org/10.15468/ksjwmn
Guarneri I, Tagliapietra D (2024). <i>Teredo bartschi</i> collection from CNR ISMAR of Venice	https://doi.org/10.15468/rfthq6

7. References

Armeli Minicante S., Birello G., Ceregato A., Perin A. (2017). Archivio Studi Adriatici (ASA) al servizio della ricerca: istruzioni per l'uso. Rapporto Tecnico IRCrES-CNR, n. 4/2017. <http://dx.doi.org/10.23760/2421-5562.2017.004>

- Armeli Minicante S., Camatti E., Capotondi L., D'Onofrio R., Di Russo E., Ferraro L., Giordano L., Grande V., Guarneri I., Maggiore F., Redolfi Bristol S., Sigovini M. (2024). L'Infrastruttura di Ricerca DiSSCo (Distributed System of Scientific Collections): il ruolo del CNR-ISMAR e le prospettive di sviluppo. Rapporto Tecnico ISMAR, n.27/2024, ISSN (ON LINE): 2611-4070. <https://doi.org/10.26383/CNR-ISMAR.2024.3>
- Armeli Minicante S., Di Russo E., Lopez-Bautista J. (2024). Species of *Ulva* (Chlorophyta) from the Lagoon of Venice (Italy). Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/ksjwmn> accessed via GBIF.org on 2024-12-12.
- Camatti E., Di Russo E., Nurra N. (2024). Species of Zooplankton from the LTER-Italy Site in the Lagoon of Venice and Oceanographic Platform CNR (North Adriatic Sea). Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/8scg2m> accessed via GBIF.org on 2024-12-12.
- D'Onofrio R., Capotondi L. (2024). North Adriatic Foraminifera Collection (NAdFC). Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/2b5bha> accessed via GBIF.org on 2024-12-12.
- Dillen M., Abraham L., Bogaerts A., De Smedt S., Engledow H., Leliaert F., Trekels M., Dessein S., Groom Q. (2024). The Meise Botanic Garden Herbarium Data Management Plan. Research Ideas and Outcomes 10: e124288. <https://doi.org/10.3897/rio.10.e124288>
- Giordano L., Ferraro L. (2024). LiCSiC - Foraminifera Linosa Sicily Channel Collection. Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/gv8k7a> accessed via GBIF.org on 2024-12-12.
- Greeff M., Kolbmann W., Burri T., Stöckli E., Klaassen S., Beck A., Chervet N., Tschudin P., Menkveld-Gfeller U., Frick H. (2021). Databasing, digitisation and data storage. In: Frick H, Greeff M (eds). Handbook on natural history collections management – A collaborative Swiss perspective. Swiss Academies Communications 16 (2). <https://doi.org/10.5281/zenodo.4316839>
- Guarneri I., Tagliapietra D. (2024). *Teredo bartschi* collection from CNR ISMAR of Venice. Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/rftqh6> accessed via GBIF.org on 2024-12-12.
- Haston E.M., Hardisty A. (2020). An Introduction to the Minimum Information about a Digital Specimen (MIDS) Digitisation Standard. Biodiversity Information Science and Standards 4: e59214. <https://doi.org/10.3897/biss.4.59214>
- Haston E., Hardisty A.R., Addink W., Dillen M., Glöckler F., Groom Q., Paul D., Petersen M., Saarenmaa H., Güntsch A., Norton B., Penzlin A., Weiland C., Rindal E., Triebel D., Walcott R., Chapman C. (2023). The Minimum Information about a Digital Specimen (MIDS) specification for Natural Science Collections (draft). V0.17. 2023-07-13
- Maggiore F., Di Russo E., Armeli Minicante S. (2024). Crustacean collection from CNR ISMAR of Venice. Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/5hgi7b> accessed via GBIF.org on 2024-12-12.
- Nanglu K., de Carle D., Cullen T.M., Anderson E.B., Arif S., Castañeda R.A., Chang L. M., Iwama R.E., Fellin E., Manglicmot R.C., Massey M.D., Astudillo-Clavijo V. (2023). The nature of science: The fundamental role of natural history in ecology, evolution, conservation, and education. Ecology and Evolution, 13, e10621. <https://doi.org/10.1002/ece3.10621>

Nieva de la Hidalga A., Rosin P.L., Sun X., Bogaerts A., De Meeter N., De Smedt S., Strack van Schijndel M., Van Wambeke P., Groom Q. (2020). Designing an Herbarium Digitisation Workflow with Built-In Image Quality Management. *Biodiversity Data Journal* 8: e47051. <https://doi.org/10.3897/BDJ.8.e47051>

Phillips S., Green L., Weech M.H. (2014). Review of digitisation workflows and equipment. <https://www.synthesys.info/content/dam/nhm-synthesys/synthesys-deliverables/synthesys3/NA3%20Del.%203.3%20Review%20of%20Digitisation%20workflows%20and%20Equipment.pdf>

Sigovini M., Keppel E., Tagliapietra D. (2016). Open Nomenclature in the biodiversity era. *Methods in Ecology and Evolution* 7(10):1217-1225. <https://doi.org/10.1111/2041-210X.12594>

Sigovini M., Guarneri I., Tagliapietra D., Sabino A., Di Russo E. (2024). NAdEm - Northern Adriatic Epimegabenthos. Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine. Occurrence dataset <https://doi.org/10.15468/r8f2c2> accessed via GBIF.org on 2024-12-12.

Wilkinson M., Dumontier M., Aalbersberg I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

The CNR-ISMAR DiSSCo Working Group



Lucilla Capotondi
Foraminifera biodiversity
Research Branch: Bologna



Simona Armeli Minicante
GBIF technical point
Seaweed biodiversity
Research Branch: Venice



Ilaria Conese
Foraminifera biodiversity
Research branch: Bologna



Elisa Camatti
Zooplankton biodiversity
Research Branch: Venice



Roberta D'Onofrio
Foraminifera collection curator
Research Branch: Bologna



Edoardo Di Russo
Collections curator
Research Branch: Venice



Valentina Grande
Storage & data manager
Research Branch: Bologna



Irene Guarneri
ISMAR representative for DiSSCo
Zoobenthos biodiversity
Research Branch: Venice



Luciana Ferraro
Foraminifera biodiversity
Research Branch: Naples



Francesca Maggiore
Zoobenthos biodiversity
Research Branch: Venice



Laura Giordano
ISMAR representative for DiSSCo
Bioinformatics & Imaging
Research Branch: Naples



Marco Sigovini
Zoobenthos biodiversity
Research Branch: Venice