

---

## EA-CoP Data Access and Sharing Policies

### EA-CoP Data Access and Sharing Policies

1. Purpose of this document
  - 1.1 Revision History
2. Introduction
3. Concepts and definitions
  - 3.1 Target for Data Access and Sharing Policies
  - 3.2 iMarine & D4Science infrastructure
  - 3.3 What are users?
  - 3.4 What type of data?
  - 3.5 Metadata
    - 3.5.1 Dynamic data
  - 3.6 Public data and shared data
  - 3.7 Copyright and license
4. Data access
  - 4.1 Open format
5. Sharing data
  - 5.1 Which license?
    - 5.1.1 Examples of possible data access and data sharing issues
  - 5.2 Quality assurance in iMarine
  - 5.3 Derivative works
6. Use cases
  - 6.1 Code list management
  - 6.2 Spread (Statistical and Geospatial data)
  - 6.3 Fisheries Linked Open Data (FLOD) , using an external knowledge base
  - 6.4 VME Reporting tool
7. Policy formulation
8. Sources of information
- Annexes
  - Annex 1 - Possible TOC and structure for an iMarine partnership-oriented policy document on data access and sharing
  - Annex 2 - Draft iMarine Terms of Use
  - Annex 3 - Draft iMarine Privacy policy
  - Annex 4 - Key principles for discussion

<b>Revision history</b>		
<b>Date</b>	<b>Author</b>	<b>Summary</b>
19 March 2012	M. Taconet, A. Gentile, A. Ellenbroek	First creation of the document
26 September 2012	M. Taconet, A. Gentile, A. Ellenbroek, F. Simeoni	<ul style="list-style-type: none"><li>- New draft version with updates based on feedback collected during the first iMarine Board Meeting (20-22 March 2012 FAO, Rome, Italy).</li><li>- Addition of par. 3.6 and par. 7</li><li>- Addition of Annexes 2 and 3 for draft "iMarine Terms of Use" and draft "iMarine Privacy policy" respectively.</li><li>- Addition of Annex 4 Key principles for discussion</li></ul>

## 1. PURPOSE OF THIS DOCUMENT

This document has been elaborated following guidance of the iMarine Board kick-off meeting, and constitutes a first attempt to compile principles, concepts, and policy elements regarding data access and data sharing in iMarine. It is meant to trigger reactions from iMarine Board members and stimulate debate towards enhanced scoping of the developing iMarine data policy.

### 1.1 Revision History

**March 2012**, a first version of the document was provided to the iMarine Board members.

Action expected from iMarine Board members: review this document prior to the first iMarine Board meeting (20-22 March 2012), and be ready to provide a general opinion on the document as well as specific points for discussion during the meeting. The meeting will strive to provide guidance for the overall structure of the iMarine data policy, as well as identify Topics which require priority developments during the next intersession.

**September 2012**, a new draft version based on feedback collected during the first iMarine Board Meeting was compiled giving priority to those topics discussed during the meeting. Drafts for “iMarine Terms of Use” and “iMarine Privacy policy” have been added in the annex.

**Action expected from iMarine Board members:** review this document during the next intersessional period and provide comments on the overall structure and content of the iMarine data policy. The annex 4 “Key principles for discussion” tries to guide the revision of the document by highlighting some relevant aspects. An agreed structure for the Policy documents will lead to the finalization of the policy document(s).

---

## 2. INTRODUCTION

The iMarine project "Description of Work" in WP3 - contains Task 3.2 Governance Mechanism, Policies and Guidelines Development:

*“The sustainable exploitation of the iMarine data e-infrastructure requires that the Ecosystem Approach Community of Practice (EA-CoP) is involved in the development of Governance mechanisms and Policies which will drive its operation, and that guidelines be available in support of its implementation. The guidelines should stipulate how the EA-CoP can benefit from the data infrastructure services”.*

iMarine provides an e-infrastructure to Communities of Practice. The “open data” infrastructure is expected to enhance science-based data processes in support to implementation of the Ecosystem Approach.

Data enter the infrastructure from different sources by the community and are then utilized to deliver new outputs. As governing instruments, data access and sharing Policies are meant to set principles and boundaries which shall govern the in-and-out flow of data by the EA-CoP, data security and confidentiality aspects, collaborative work through data sharing, and publishing.

This document has to be understood as a “work-in-progress”. It doesn’t necessarily reflect the final structure of the Policy document (proposals for such structure are found in annex) and rather strives to bring the iMarine Board to a common understanding and agreement on concepts, principles and boundaries.

---

### 3. CONCEPTS AND DEFINITIONS

This paragraph provides a set of definitions required for smooth understanding of concepts as used in the context of iMarine.

#### 3.1 Target for Data Access and Sharing Policies

Policies are needed to:

(data access)

- guarantee the integrity of all data populating iMarine
- facilitate the access and the utilization of data in cost effective ways

(data sharing)

- manage private and confidential data
- share data among users according agreed principles
- properly inform users on data, their provenance and conditions of use, including:
  - propagate copyright and license settings
  - manage the document metadata
  - provide documentation of sources

#### 3.2 iMarine & D4Science infrastructure

iMarine utilizes the distributed infrastructure of D4Science (D4S - Distributed collAboratories Infrastructure on Grid ENabled Technology 4 Science) to provide computational resources to a target community of users. iMarine resources include hardware, datasets, and software services. The *datasets* are the input and output of those processes, hence the core digital assets of the community. The *services* capture the processes that are directly or indirectly relevant to the needs of the target community. *Hardware* resources is where processes unfold and datasets are stored. It should be noted here that sharing and harmonization of software and hardware resources should be matter of another Policy document.

The enabling technology for D4Science is gCube, the software system for building e-Infrastructures for the integration and management of computational resources at limited or no cost for the target community.

gCube allocates resources to distinguished groups of users. Long-lived resource allocations for broad user groups are known as Virtual Organisations, or VOs. Resources of direct relevance to the iMarine community are allocated to a specific D4Science VO. More dynamic and task-oriented allocations of VO resources to subsets of VO users are referred to as Virtual Research Environments, or VREs.

The system also defines interactive interfaces to VOs and VREs for both end users and administrators. VREs, in particular, are presented to their users through collaborative environments hosted in an Infrastructure-wide portal, the iMarine Gateway. Through these environments, scientists and practitioners can access cross-disciplinary dataset and knowledge, as well as a rich array of services that enable innovative analysis, visualization and domain specific knowledge generation processes. In a VRE, users can access data managed by the D4S infrastructure, but also data available from external resources which are public, e.g. SDMX repositories, Dublin Core libraries, GeoServer, GeoNetwork, Fisheries Linked Open Data (FLOD).

VREs can only access data from other users in other VREs where the latter have explicitly granted access, regardless of the license type. External public data can be used according to their accessibility. D4Science does not manage external license schemes (see terms of use).

### 3.3 What are users?

Any agent (human or software) granted to access iMarine data (through the D4Science infrastructure) is a user. Data providers and data consumers are here considered as users.

From the infrastructure perspective, users own all data they upload.

### 3.4 What type of data?

Data are clustered in “data collections”, intended as an ordered group of elements, all of the same type, e.g. a collection of aquatic species distribution maps or collections of global fishery productions statistics. Each data collection can contain one or more dataset.

iMarine datasets vary in semantics, structure, format, access and storage technology, and may include:

- collections of structured data exposed by dedicated access services within the infrastructure, through collections of documents and metadata descriptions thereof, to databases (e.g. geospatial databases, taxonomic databases, statistical databases);
- structured or unstructured data files passed in input to, or produced in output by, interactive or batch processes executed on the infrastructure (e.g. a .csv file used to import a statistical dataset, a report instantiated within a VRE from a template), or else uploaded by users and shared among users in VREs (e.g. documents, images, video).

The data coverage extends to any recorded information regardless of the format. The policy here considered is applied to any of the above dataset type and users must be made aware of the specific policy that applies to the related data collection.

### 3.5 Metadata

All datasets in the D4Science infrastructure are accompanied by a minimum set of metadata.

To develop a consistent access and sharing policy, all datasets submitted to iMarine are to be associated with a descriptive and standard set of metadata, the business metadata (i.e. ownership and context, authorship, copyright licenses, content description and main features).

The business metadata, defined by the data provider, are not supposed to be changed by other users and notwithstanding other purposes, can be utilized to give instructions on how to access and share the various collections of data. Format of metadata may change according to the format of the dataset.

Here follows a list of candidates business metadata (applied at dataset level):

#### Ownership and context

- Owner
- Context (*including Spatial scale, geographic coverage, topic coverage, and references in case of derivative works*)

#### Authorship

- Author
- Title
- Publisher
- Publish date
- Last update date
- Expiry date

#### Copyright licenses

- Rights management, Creative Commons License type (*or other licenses ...*)

### Content description

- Data aggregation level
- Spatial Scale
- Geographical coverage
- Topic coverage
- ...

Metadata already associated to a dataset must be preserved during the upload process and in case further integrated, i.e. by adding the context. It should be considered that the only categories of works for which CC does not recommend its licenses are computer software and works that are no longer protected by copyright or are otherwise in the public domain.

#### **3.5.1 Dynamic data**

The iMarine data policy primarily focuses at the level of datasets.

The data contained in a dataset are too varied to apply a single policy to.

At application level, for certain data types, individual policies can be defined, however these are not iMarine policies. An example is SDMX, where the structure allows for the definition and tracking of metadata at observation and time series level. These should be regarded as SDMX policies.

iMarine can offer support to business metadata management but at this stage is not yet defined how and with which level of detail and automatism; i.e. data points may change dynamically along time and the metadata should reflect such changes to warn users when producing new outputs.

#### **3.6 Public data and shared data**

Data are public when exposed to general view with no restrictions (except for possible registration procedures), when openly accessible. When users in iMarine publish a dataset means that the data are exposed to the general public (e.g. through data.i-Marine.eu).

Exposing data within the iMarine infrastructure (i.e. sharing through one or more VREs) is a different action where data are only available under restricted access.



### 3.7 Copyright and license

Here follows some definitions for better understanding the meaning and the relation between a copyright and a license (or set of licenses).

*[Copyright is a form of protection provided for original works of authorship, including literary, dramatic, musical, graphic and audiovisual creations. "Copyright" literally means the right to copy, but has come to mean that body of exclusive rights granted by law to copyright owners for protection of their work.*

Source: [YouTube](#)]

*[Creative Commons licenses are copyright licenses – plain and simple. Period. CC licenses are legal tools that creators can use to offer certain usage rights to the public, while reserving other rights. Without copyright, these tools don't work. Source: [Creative Commons](#)]*

*[A **copyright safeguards the ownership of an intellectual property**. If you hold copyright to some intellectual property, you have several rights regarding that property, and you can assign (sell or give) some or all of those rights to others. **A License**, on the other hand, **is a document** lets someone use your intellectual property. Source: [LinuxJournal.com](#)]*

*[Copyright law gives the creator of a work the exclusive right to copy, alter, distribute, or perform/display the work, **or authorize others to do so**. A Creative Commons license **is that authorization**. Source: [Answers.com](#)]*

---

## 4. DATA ACCESS

In iMarine, data are available through VREs, hence the first way to allow the access to a given collection of data is granting those users to access the desired VRE(s). The administrator of the VRE is enabled to grant access rights to users. Different levels of user authorization are envisaged such as manager, publisher, editor, viewer, *etc.*

By accessing the data stored in iMarine, users agree to the related Data Access Policy.

Public data can be accessed by the general public through a dedicated website, e.g. “data.i-Marine.eu” and through dedicated web applications (e.g. web services).

### 4.1 Open format

*An open format is one that is platform independent, machine readable, and made available to the public without restrictions that would impede the re-use of that information [Source: Open Government Directive of December 8, 2009.]*

iMarine is importing and exposing a variety of data types in open format. The list of Open formats handled in iMarine is agreed upon aware of the functional needs, costs and benefits.

## 5. SHARING DATA

*In most jurisdictions there are intellectual property rights in data that prevent third-parties from using, reusing and redistributing data without explicit permission. Even in places where the existence of rights is uncertain, it is important to apply a license simply for the sake of clarity. Thus, if you are planning to make your data available you should put a license on it — and if you want your data to be open this is even more important. [Source: Guide to Open Data Licensing - Open Knowledge Foundation <http://opendefinition.org/guide/data/>]*

### 5.1 Which license?

*A license is an agreement through which a licensee leases the rights to a legally protected piece of intellectual property from a licensor – the entity which owns or represents the property — for use in conjunction with a product or service. [Source: <http://www.licensing.org/education/introduction-to-licensing/> ]*

Data loaded in the D4Science infrastructure are covered by a license type; iMarine is relying upon the **Creative Commons** copyright licenses.

By default, data uploaded by users remain in the private domain under a restricted license (e.g. CC BY-NC-ND).

The license type of a dataset can be changed by modifying the proper metadata field (e.g. switching to CC BY or CC BY-NC).

*The “CC BY” license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.*

*The “CC BY-NC” license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don’t have to license their derivative works on the same terms. [Source: Creative Commons copyright licenses <http://creativecommons.org/licenses/>]*

The adoption of such license type implies that in any derivative work the source of information must be always acknowledged. A detailed quality assurance report (can also be referred to as “Metadata for ownership and context”) may be compiled for the derivative works to better describe how the shared data were utilized (e.g. indicating the coverage, data structure, sources).

---

### 5.1.1 Examples of possible data access and data sharing issues

VTI or VME records

The provider requires special credentials

Only Staff members of the RFB can access occurrences

The provider has a formatted data source (e.g. .xls)

iMarine may need to support the format NAF

The provider requires access protocol

The owner provides data covered by a (restricted) license or policy.

The owner only can share aggregated data under a public license

The owner has to keep occurrence data private

There are no copyrights claimed on public data (the aggregated data)

Minimal requirements for aggregation are Month, .5 degree, and at least 20 occurrences.

### 5.2 Quality assurance in iMarine

Each data collection accessible through iMarine is documented with its associated policy, called “data collection policy”, it enables users to evaluate the information content and conditions for utilization. Data owners are providers and responsible for the maintenance of such documentation.

A data collection policy should include mandatory and optional fields.

Mandatory: business meta data and key quality criteria (e.g. sources, timeliness).

Optional: any other element enabling users to evaluate the quality of the dataset.

All elements of a data collection inherit the properties of the data collection and can be cited. The citation includes reference to the data collection policy.

Also single files uploaded in iMarine should be subject to a data collection policy, in particular as soon as such file is to be shared with other users, and/or integrated for derivative work. The process of developing a data collection policy to such single files should be a simple one (as far as possible), e.g. through merging some basic compulsory Metadata for the file with default iMarine Data collection policy document.

The data collection policy also procures a relevant guideline for any possible derivative work.

---

In the FAO Fisheries and Aquaculture Department (FI) and in the Fishery Resources Monitoring System (FIRMS) similar documentations are called “Data collection” and “Data quality assurance”.

Examples:

FI Data collection fact sheets <http://www.fao.org/fishery/statistics/collections/en>

FIRMS Data quality assurance fact sheets <http://firms.fao.org/firms/contributions/en>

### 5.3 Derivative works

As stated in Data.Gov (US Government):

#### ***Secondary Use***

*Data accessed through Data.gov do not, and should not, include controls over its end use. However, as the data owner or authoritative source for the data, the submitting Department or Agency must retain version control of datasets accessed. Once the data have been downloaded from the agency's site, the government cannot vouch for their quality and timeliness. Furthermore, the US Government cannot vouch for any analyses conducted with data retrieved from Data.gov.*

iMarine may adopt a similar policy although mechanisms and tools for updating data within the infrastructure vs. original sources are foreseen.

The business metadata, including the license type, give all set of instructions for data utilization. When combining different data sources, the more restricted source of information drives the overall possibility of public dissemination.

A sort of cascading mechanism, if only one of the sources is confidential, any output can be only confidential!

## 6. USE CASES

To better understand the need of data access and sharing policies, a set of use cases describe likely cases where such policies are needed. The list is not complete, but may serve as a guide for further development.

### 6.1 Code list management

*- Copy or refer to a (subset) of other people's code lists and re-use it -*

Code list in iMarine are supposed to be public; they can be accessed by all users, and freely shared for further utilization.

However, this clearly leaves many gaps where data have to be managed; the code lists may be too short (no code for 'Unknown'), or the adverse; too long (1.4M species codes). Or the code lists may have versions over time, space, or user-groups.

This requires a code list management in iMarine.

A confidential set of codes in a given VRE is not foreseen to be utilized for public use.

The use-case for code-list management thus becomes:

A user identifies a code list, and reviews the attached policy;

If the user needs to modify the policy, a copy must be created;

The user then redefines the policy (e.g. the validity over space and time for a domain) and re-publishes the code-list in the public domain.

### 6.2 Spread (Statistical and Geospatial data)

*- Change user data and republish; how to deal with data owners -*

Spread manipulation can be done according to the original dataset license scheme. Data originator should be mentioned but no responsibility can be assigned to such owners. The SPREAD data access and sharing use-case thus becomes:

A user re-allocates a time series over its spatial dimension;

The result, in itself a time series object, is stored in the SPREAD VRE, and is only accessible by the user that created it. This TS can be shared through the VRE workspace;

---

The authorized users reviews the results, and may decide to publish the results as a private geospatial product, visualized on a map-interface;

The new product can be kept as a private map, or published in a open repository (e.g. data.i-Marine.eu), where the product can be discovered and accessed through the metadata, or directly through web applications.

### **6.3 Fisheries Linked Open Data (FLOD) , using an external knowledge base**

*- How to mix Open Data with proprietary data; how to manage attributions? -*

Each dataset has its own license scheme, the proprietary data can be mixed with open data according to their license scheme (e.g. a dataset “CC BY” can be used with no restrictions, a dataset “CC BY-ND” can be copied distributed but no derivative works are allowed). Attributions must be given to all dataset utilized. In case of FLOD, the application should provide all relevant sources of information.

The FLOD use case cannot be described as of today.

### **6.4 VME Reporting tool**

*- How to ensure confidentiality -*

VME data are stored in the Virtual Research Environment (VRE) which is a collaborative environment where data owner grant selected users to access the data and to work on them. The business metadata (including the license scheme) give information on how such data can be utilized and disseminated. Authorized users manipulate the data according to the license scheme (e.g. by aggregating) and make them available for publishing . All published data are no longer subject to any confidential issue.

The VME data access and sharing use-case thus becomes:

The user generated a report on a VME in the VME-VRE. The report is private.

If the user needs to share the report, a user can be selected from this or any other VRE.

Alternatively, the user can use the Work-flow facility to forward a document to a next stage.

Confidential data are not managed through the Work-flow; if a user includes confidential data in a report, other users will be able to access that information.

If a report is ready to be published, the assigned persons in the Work-flow can publish the report in a variety of formats, including XML FIMES compliant.

Until publishing, it is not visible, discoverable or accessible to users beyond the VRE-boundaries.

## 7. POLICY FORMULATION

Most Internet service providers inform on their policies through web pages and specific documentation. Such policies may include terms of services, data privacy, copyright and legal aspects. These documents describe how services are delivered, how these are best utilized and cover the relation between providers and users.

In the iMarine context, some of these policies can be considered too strict, but clear and unequivocal policies contribute to maintain a high level of community trust and data integrity.

Each service provider adopts different solutions to describe its policies, some examples are included in Table 1. The most common practice seems to provide information under two main category pages: “Terms of use” and “Privacy”. In addition, for core activities, dedicated pages on specific policies and copyright are usually included (e.g. YouTube, FAO) or Data.Gov.Uk who also created specific pages for users’ code of conduct.

“Terms of Service”, “Terms of Use” or equivalent terminology can cover a range of issues, including acceptable user behavior online, a site dissemination policy, and data citation notices.

Most organizations reserve the right to restrict a user's access to the service if they violate the terms in the agreement. In serious cases, the user may have his or her account terminated.

**Table 1.** Examples of types of policy pages

Website	Terms of Use	Policy	Copyright	Disclaimer	Privacy
<a href="#">Data.Gov.Uk</a>	Terms and conditions	Code of Conduct ; Moderation Policy			X
<a href="#">Data.Gov</a>		Data policy			Privacy policy
<a href="#">CC</a>	X	Policies			X
<a href="#">YouTube</a>	X	Safety	X		X
<a href="#">Google</a>	X				X
<a href="#">UN</a>	X		X	Fraud Alert	Privacy Notice
<a href="#">FAO</a>			X	Scam Alert	Privacy policy



<a href="#">EU - Eurostat</a>			X	X	Personal data protection ; Cookies
<a href="#">NSDL</a>	X				X
<a href="#">iMarine</a>				X	Data Privacy
<a href="#">AOL</a>	X	AOL member community guidelines			Privacy policy
<a href="#">FaceBook</a>	Statement of Rights and Responsibilities	Community Standards			Data Use Policy
<a href="#">OBIS</a>	Web page citations; Quality Control				
<a href="#">CoL</a>			X	X	
<a href="#">GBIF</a>	<a href="#">GBIF Data Use Agreement</a> ; <a href="#">GBIF Data Sharing Agreement</a> ; <a href="#">Memorandum of Understanding</a>				
<a href="#">ITIS</a>	Website Citation; Submit Data (Requirements and Guidelines)				Privacy Statement and disclaimer
<a href="#">MyOcean</a>	<a href="#">SERVICE COMMITMENTS AND LICENCE</a>		All rights reserved		
<a href="#">Geonetwork</a>			X		
<a href="#">ESA</a>	Portal terms of use - including copyright		X		

## 8. SOURCES OF INFORMATION

For the present discussion paper, the following policies have been consulted:

Creative Commons copyright licenses <http://creativecommons.org/licenses/>

CC Wizard for choosing license <http://creativecommons.org/choose/>

CC FAQ <http://wiki.creativecommons.org/FAQ>

YouTube Copyright Center [http://www.youtube.com/t/copyright\\_center](http://www.youtube.com/t/copyright_center)

YouTube & CC [http://www.youtube.com/t/creative\\_commons](http://www.youtube.com/t/creative_commons)

Data.Gov <http://www.data.gov/datapolicy/>

Example of access interface to a Data.Gov dataset <http://explore.data.gov/d/u9mw-zn8r>

Data.Gov.UK <http://data.gov.uk/terms-and-conditions> (Open Government Licence)

Aquatic Sciences and Fisheries Abstracts (ASFA) [ftp://ftp.fao.org/FI/asfa/faq/faq10\\_e.pdf](ftp://ftp.fao.org/FI/asfa/faq/faq10_e.pdf)

FAO procedures for data licensing (informal document kindly provided by Ms R. Tucker)

GBIF Data Sharing Agreement <http://data.gbif.org/tutorial/datasharingagreement>

FishFrame <http://www.ices.dk/products/CMdocs/2006/M/M0206.pdf>

National Science Digital Library (NSDL) <http://nsdl.org/about/policies-and-guidelines>

UN Media <http://www.unmultimedia.org/photo/guidelines.jsp> (and other relevant pages)

FAO Aquaculture Photo Library <http://www.fao.org/fishery/photolibrary/home/en/>

SeaDataNet <http://www.seadatanet.org/Media/seadatanet/Files/Publications/SeaDataNet-Data-Policy>

Metadata

Wikipedia: <http://en.wikipedia.org/wiki/Metadata> ;  
[http://en.wikipedia.org/wiki/Preservation\\_Metadata](http://en.wikipedia.org/wiki/Preservation_Metadata)

Dublin Core Metadata Initiative Wiki  
[http://wiki.dublincore.org/index.php/Glossary/Resource\\_Discovery](http://wiki.dublincore.org/index.php/Glossary/Resource_Discovery)

FAO Data blog <http://faodata.blogspot.com>

gCube Wiki [https://gcube.wiki.gcube-system.org/gcube/index.php/GCube\\_Wiki](https://gcube.wiki.gcube-system.org/gcube/index.php/GCube_Wiki)

D4Science privacy policy <http://www.d4science.org/privacy>

iMarine privacy policy <http://www.i-marine.eu/Pages/DataPrivacy.aspx>

iMarine Disclaimer / Terms of use <http://www.i-marine.eu/Pages/Disclaimer.aspx>

gCube [https://gcube.wiki.gcube-system.org/gcube/index.php/GCube\\_Wiki](https://gcube.wiki.gcube-system.org/gcube/index.php/GCube_Wiki)

Vocabulary of Interlinked Datasets (VOID) <http://vocab.deri.ie/void> (RDF Schema vocabulary for expressing metadata about RDF datasets.)

GeoRepository, advanced authorization manager for GeoServer <http://geo-solutions.blogspot.com/2011/05/preview-georepository-advanced.html>

---

## Annexes

### ANNEX 1 - POSSIBLE TOC AND STRUCTURE FOR AN IMARINE PARTNERSHIP-ORIENTED POLICY DOCUMENT ON DATA ACCESS AND SHARING

#### D4Science platform

- Security policies

- Basic access principles (e.g. registered users)

- Primary reference to EU standards (e.g. INSPIRE directive, Data directive on Privacy)

#### iMarine partnership

- Definitions (Open data, open access, license, open format, publishing, ...)

- Principles

  - Inherit D4Science platform policies

  - Commitment to Open access and Open data

  - A Partnership working towards delivery of authoritative source of information in support to EA

  - Other principles

- Open access

  - Easy to retrieve

  - Usable (e.g. GIS shape files have to meet some quality requirements)

  - Open formats supported

- Open data (Open to be understood as available under existing data policies)

  - Citation and credit rules (on the ethical sense)

  - Primary responsibility to data owner to define the applicable policy

    - Highlighting scope or range of possibilities

  - Positioning confidentiality

- License (all data must be licensed)

Default license

Eligible license

iMarine license ?

Quality assurance in iMarine

Principles

Authoritative source of EA information

Coherence, non duplication, proper documentation, ...

Quality of sources: Reliable, updated,

Meaningful, through explicit semantics

Data.i-Marine.eu for disseminating iMarine quality assured products

Business Metadata

Data collections and their policies

Version control

Provenance of derived products

Principle of traceability

Stress responsibility of analyst in the products elaborated (e.g. if uses own data without proper Metadata)

Standard schemas supported

Reference lists supported

---

## ANNEX 2 - DRAFT IMARINE TERMS OF USE

The iMarine Terms of Use consists of this Chapter, the iMarine Privacy Policy, and the iMarine Community Best Practices.

### Disclaimer

The iMarine Gateway has been produced with the co-funding of the European Commission. The content of the Gateway is the sole responsibility of the iMarine consortium and cannot be considered to reflect the views of the European Commission.

### Copyright

The data and information available through <http://www.i-marine.eu> are available under terms described in the "license" or "constraints" field of individual dataset records (meta-data). Except where otherwise noted this is the **Creative Commons License**.

All dataset records (meta-data) published on [www.i-marine.eu](http://www.i-marine.eu) are licensed under the [Creative Commons License](#).

By accessing the data stored in iMarine, users agree to the related Terms of Use.

### Your Agreement to the Terms

YOUR ACCESS OR USE OF ANY WEBSITE, VRE OR SERVICE IN ANY WAY SIGNIFIES THAT YOU HAVE READ, UNDERSTAND AND AGREE TO BE BOUND BY THE TERMS. By accessing or using any Website, VRE or Service you also declare that you have the legal authority to accept the Terms on behalf of yourself and any party you represent in connection with your use of any Website, VRE or Service. If you do not agree to the Terms, you are not authorized to use any Website, VRE or Service.

### iMarine On-line Community

The on-line community of iMarine is composed of a support community organized in the iMarine **Channel**, and the active users of the data and applications through the iMarine **Gateway**.

Users can register themselves as a "iMarine **Channel** Community Member" to request access to services provided by iMarine such as newsletter, download area, video repository, personal web space.

Please consult the [Privacy](#) page and [iMarine Community Best Practices](#).

In addition, users can register themselves on the iMarine **Gateway** to join one or more Virtual Research Environments (VRE's) and / or gCubeApps. This request will have to be approved by the VRE manager.

### iMarine Gateway

The iMarine gateway is the access point to iMarine products and services.

---

iMarine is welcoming any institution in the field of fisheries management and conservation of marine living resources.

Access is granted to individual users who are members of those institutions/organizations having an established agreement with the iMarine consortium.

### **Virtual Research Environments (VREs)**

In iMarine, collaboration services are available through Virtual Research Environments (VREs).

Only authorized users access data and services exposed through these VREs.

There are different levels of user authorization available: VRE manager, VRE User.

VRE managers have the right to grant access to other users and to assign roles.

Users can upload and manage data within dedicated VRE(s) according to their authorization level.

VRE's can be created for a specific period and a specific task, and the VRE Designer or the VRE Manager selects the services that are made available in the VRE. A document defining the roles at iMarine e-Infrastructure level is available at [http://wiki.i-marine.eu/index.php/Virtual\\_Research\\_Environments\\_Deployment\\_and\\_Operation:\\_Procedures](http://wiki.i-marine.eu/index.php/Virtual_Research_Environments_Deployment_and_Operation:_Procedures).

By default, data uploaded in iMarine are private and under the ownership and confidentiality rules of the User of the VRE who is the custodian of the data on behalf of the owning institution or organization.

Each dataset uploaded to e-Infrastructure is accompanied by a semi-automated collection of descriptive metadata.

Only metadata formats considered as iMarine standards can be fully supported. All other metadata can not be preserved during the upload process and must be added manually, i.e. by adding a coverage type.

### **Posting Content on iMarine Services**

You can post content to a Service only if (a) you created and own the rights to the content or you have the owner's express permission to post the content; and (b) the content does not infringe any other person's or entity's rights (including the copyrights, trademarks, or privacy rights) or violate any applicable laws, this Terms of Use, our Community Guidelines, or any other posted policies. iMarine can remove content for any infringing reason.

Users are expected to post content compliant with the iMarine objectives and scope. The e-Infrastructure is supporting the principles of the "Ecosystem Approach to Fisheries Management and Conservation of Marine Living Resources" and it empowers practitioners from multiple scientific fields such as fisheries, biodiversity and ocean observation.

You are responsible for any content you post to our Services and the consequences of sharing or publishing such content with others or the general public. This includes, for example, any personal information, any confidential data or any unofficial data. iMarine is not responsible for the consequences of sharing or posting any personal or other information on its services.

When you use a service to **load** content into a VRE, you retain the irrevocable, exclusive, worldwide right and license to use, reproduce, modify, display, remix, redistribute, create derivative works, and syndicate your content in any medium and through any service. Only where you have specifically shared, or used a specialized service for sharing data, will your data become a derivative works to be licensed under a schema selected by the owner.

### **Shared Data**

iMarine offers data sharing at different levels: within a VRE, among different VREs and by publication to data repositories that may be accessible to external users. Data sharing in the e-infrastructure is possible for registered users only.

During the publication step, a CC license must be selected, metadata must be filled (if not automatically compiled) and a preferred citation must be indicated.

When you use a Service that allows users to **share, transform, readapt, modify, or combine user content with other content**, you grant us and our users an irrevocable, non-exclusive, royalty free, perpetual, worldwide right and license to [use, reproduce, modify, display, remix, perform, distribute, redistribute, adapt, promote, create derivative works, and syndicate] your content [in any medium and through any form of technology or distribution] and to permit any derivative works to be licensed under these same license terms.

Any user accessing shared data must behave according to the copyright license of the dataset. However, the e-Infrastructure does not control or audit access to and / or audit use of shared data. It can therefore not impose a shared data policy, but can encourage sharing according to best practices.

Exposing data within the iMarine infrastructure (i.e. sharing through one or more VREs) will make them accessible only to authorized users, and cannot be accessed from outside.

### **Public Data**

The VRE authorized users can publish data and make it available to the public exposing to general view with no restrictions except for a registration procedure. Each published set of data is associated to a copyright license.

Published data can be accessed by the general public through the iMarine Web Channel and through dedicated web applications (*e.g.* web services).



**Secondary Use**

Data accessed through Data.i-Marine.eu do not, and should not, include controls over its end use. However, as the data owner or authoritative source for the data, the submitting Organizations or Institutions must retain version control of datasets accessed. Once the data have been downloaded from the site, iMarine cannot vouch for their quality and timeliness. Furthermore, iMarine cannot vouch for any analyses conducted with data retrieved from [www.i-marine.eu](http://www.i-marine.eu).

**Derivative Work and Data Citation**

Derivative work should only be produced by complying with the terms and conditions established in the license of the used dataset(s).

Derivative work may contain data retrieved from the e-Infrastructure.

Every dataset should include in its metadata at least a preferred citation, a reference to the e-infrastructure source dataset and its generation date, and the date that data were accessed or retrieved from [www.i-marine.eu](http://www.i-marine.eu). If the dataset in iMarine was a composition of multiple datasets with different citations, the owner of the dataset should have added that to the relevant metadata.

If such derivative work is produced through VREs, iMarine generates a default citation, which matches references to sources used for the derivative work and their respective citations.

Finally, users must clearly state that "iMarine cannot vouch for the data or analyses derived from these data after the data have been retrieved from i-Marine.eu."

These Terms of Use are Effective as of September 15, 2012.

---

**ANNEX 3 - DRAFT IMARINE PRIVACY POLICY****iMarine Data Handling Privacy Policy & On-line Registration Privacy Statement**

Data collected through the iMarine on-line registration process is subject to Registration Process and Data Collection regulations, the European Directive 95/46 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. While registering to iMarine Channel, your personal data will be collected and further processed in order to profile your interests and ensure you receive information accordingly.

**What personal information is collected, for what purpose and through what technical means?**

The data collected is for the sole purpose of profiling membership to the iMarine Web Channel and any related messaging that may result from registration, which will be entirely within the remit of iMarine. Personal data collected include: name, surname, organisation, occupation, organisation type, sector(s), city, country, email, and area(s) of interest. Registered members will receive confirmation of their User ID and password, in order to: Access content on the iMarine Channel and Receive notification of iMarine initiatives, project news and developments.

**Who has access to your data and to whom is it disclosed?**

No personal data is transmitted to third parties nor to members of the iMarine Consortium, other than the iMarine Partnership Members. Data is not accessible by other registered members. Statistical information, such as professional role, country provenance, and organisation type, is provided to the European Commission or to iMarine Partnership, for the sole purpose of reporting on iMarine Community Membership.

**How is your data protected and safeguard?**

All the personal data collected is stored through means which must guarantee data protection and confidentiality as required by the Regulation (EC) 45/2001.

**How can personal data be verified, modified or deleted?**

The User Profile can be verified and modified at any time. To unsubscribe from the Channel, follow the steps indicated at the bottom of your User Profile page.

**How long is data retained?**

Personal data is retained for as long as the Web Channel is available and operational for its user communities. Registered members will be informed when their community or membership thereof will be terminated.

**VRE security issues**

Storage: the storage of data within the VREs is guaranteed as long as long the iMarine Gateway is available.

Access to VREs and other resources (e.g. web services) is ruled by the iMarine authentication and authorization mechanism. All data in VREs are protected through encryption mechanism and user credentials are always requested for accessing any VRE.

Confidentiality of the data is preserved as long as authorized users do not publish their data. iMarine do not change the privacy settings of any VRE without explicit request.

All messages, annotations and any other communications within a VRE or among VREs are not accessible to the public.

**Changes to this Policy**

The i-Marine.eu privacy policy will be revised or updated if practices change. You should refer to this page regularly for the latest information and the effective date of any changes. If changes are made to this policy, a new policy will be posted on our site and the date at the bottom of the page will be updated. Changes to the policy will not apply retroactively.

---

**ANNEX 4 - KEY PRINCIPLES FOR DISCUSSION**

Here follows a list of key proposed principles which should be considered and discussed by the board members for better orientating the further compilation of the policy documents.

- iMarine policy documents will consist of three web pages:
  - iMarine Terms of Use
  - iMarine Privacy Policy
  - iMarine Community Best Practices
    - Annex 1: iMarine standard formats and data exchange protocols
    - Annex 2: iMarine candidate standards
    - Annex 3: Reference guidelines
- Creative Commons is the adopted copyright licenses
  - Are there any exceptions and how should be handled?
  - Rules for derivative works - Are CC licenses enough?
- Metadata:
  - Business metadata - Which is the minimum set of metadata? Which support can iMarine offer for business metadata management?
  - Metadata are applied at dataset level not a data record level
  - Metadata format as iMarine standards:
    - principles of eligibility of a metadata format as iMarine standard
    - conditions to be met by the infrastructure for a metadata format to become an iMarine standard
  - Inheritance of the metadata in derivative works - What should be inherited, at which details level? Are envisaged any automatic mechanism?
- Definitions and concepts
  - Public and Shared data - the specificity of iMarine;
- iMarine disseminate data through a dedicated website (e.g. data.i-marine.eu)
  - Any registration procedures for accessing public data.