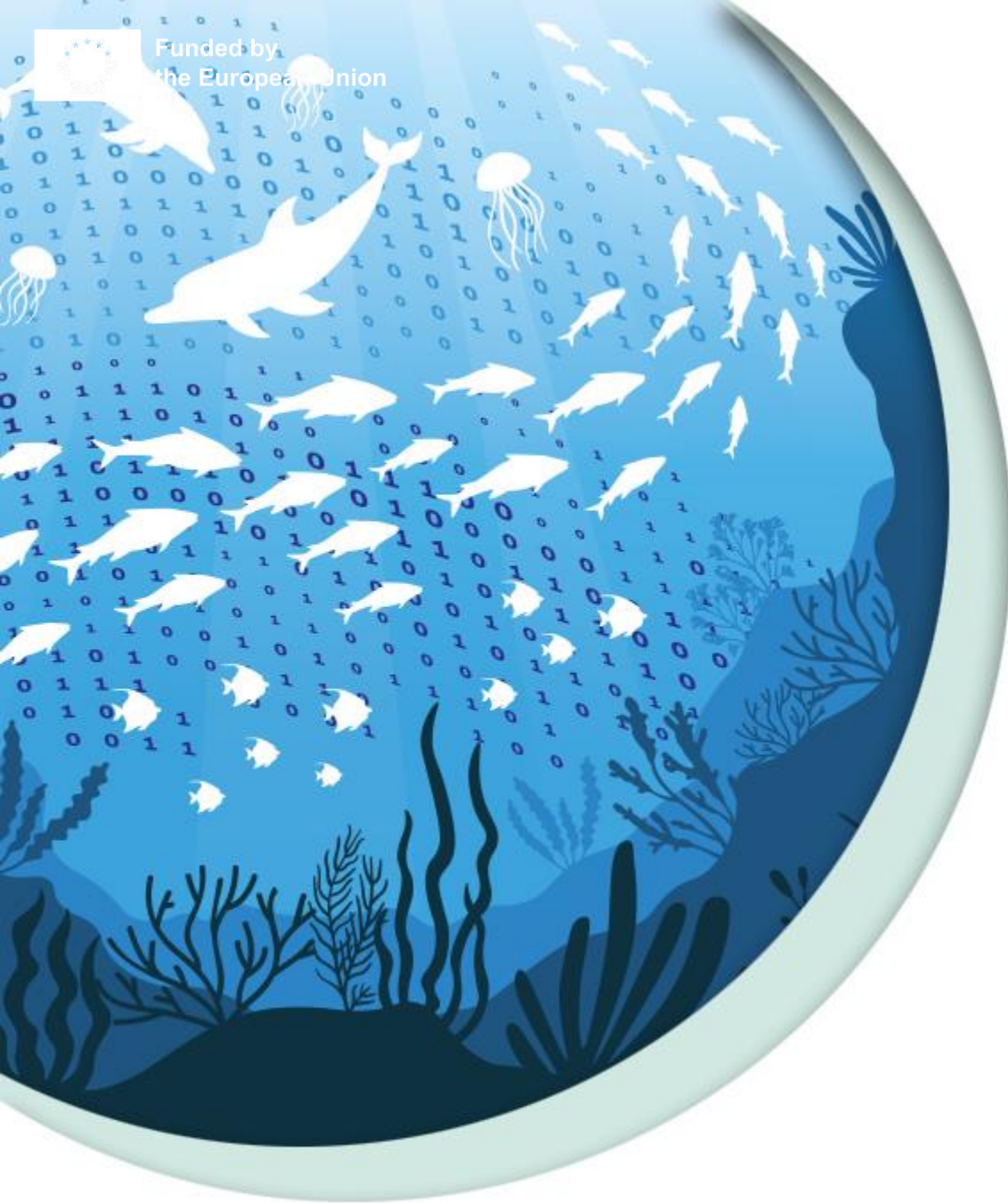




Funded by
the European Union



DTO-BioFlow

Integration of biodiversity monitoring
data into the Digital Twin Ocean

DTO-BioFlow data training
workshop:

Value standardization

Value standardization

- Consistent values not only within a dataset, but across datasets
- Values can be standardized using a controlled vocabulary, or by using a specific format or data type

language	
Identifier	http://purl.org/dc/elements/1.1/language
Definition	A language of the resource.
Comments	Recommended best practice is to use a controlled vocabulary such as RFC 5646.

Examples

en (for English)

es (for Spanish)

eventDate

Identifier	http://rs.tdwg.org/dwc/terms/eventDate
Definition	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.
Comments	Recommended best practice is to use a date that conforms to ISO 8601-1:2019 .

month

Identifier	http://rs.tdwg.org/dwc/terms/month
Definition	The integer month in which the dwc:Event occurred.
Comments	

Examples

1 (January)

10 (October)

Value standardization

- ≡ Consistent values not only within a dataset, but across datasets
- ≡ Values can be standardized using a controlled vocabulary, or by using a specific format or data type
- ≡ In some cases values can be standardized by adding an ID from a controlled vocabulary for that value in a separate column

scientificNameID

Identifier	http://rs.tdwg.org/dwc/terms/scientificNameID
Definition	An identifier for the nomenclatural (not taxonomic) details of a scientific name.
Comments	
Examples	urn:lsid:ipni.org:names:37829-1:1.3

higherGeographyID

Identifier	http://rs.tdwg.org/dwc/terms/higherGeographyID
Definition	An identifier for the geographic region within which the dcterms:Location occurred.
Comments	Recommended best practice is to use a persistent identifier from a controlled vocabulary such as the Getty Thesaurus of Geographic Names.
Examples	http://vocab.getty.edu/tgn/1002002 (Antártida e Islas del Atlántico Sur, Territorio Nacional de la Tierra del Fuego, Argentina).

Value standardization

- ≡ How to standardize depends on the field
 - ≡ General recommendations: check DwC terms definitions and comments
 - ≡ More specific recommendations: repository (e.g. OBIS, EMODnet Biology)
- ≡ Let's look into more detail into:
 - ≡ Taxonomy
 - ≡ Geography
 - ≡ Time

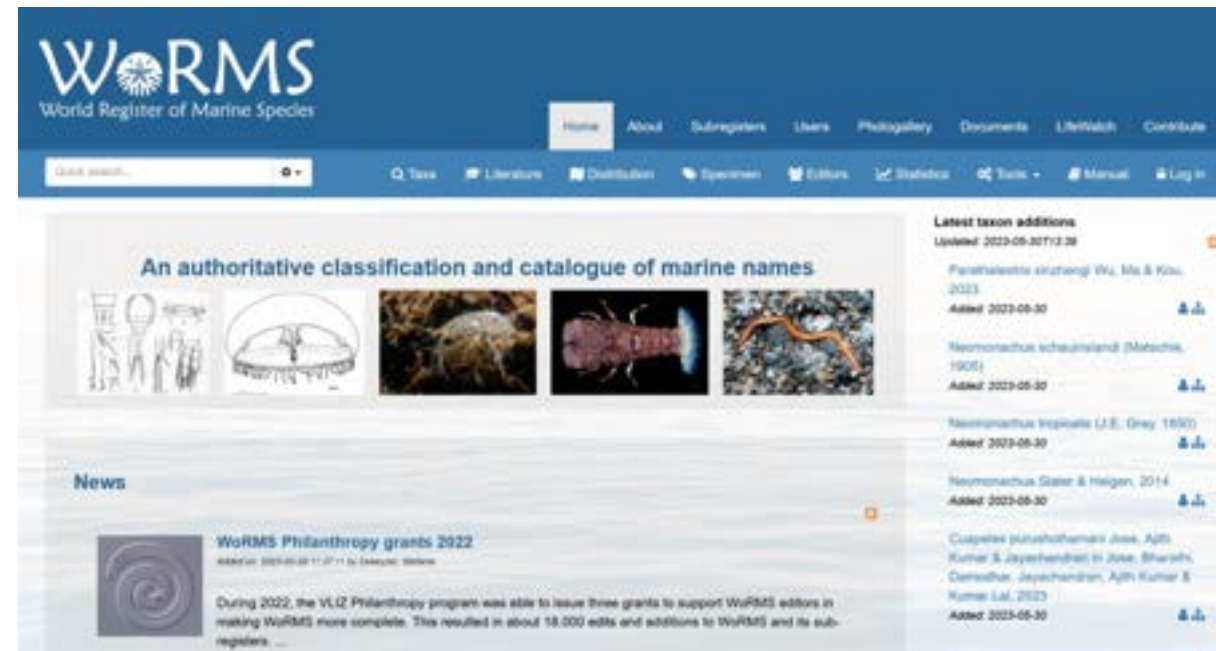
Value standardization

- ≡ How to standardize depends on the field
- ≡ Let's look into more detail into:
 - ≡ **Taxonomy**
 - ≡ Geography
 - ≡ Time

Taxonomic standardization

≡ Match names to an authoritative taxonomic register

≡ Taxonomic backbone of OBIS:
World Register of Marine Species
(WoRMS)



Taxonomic standardization

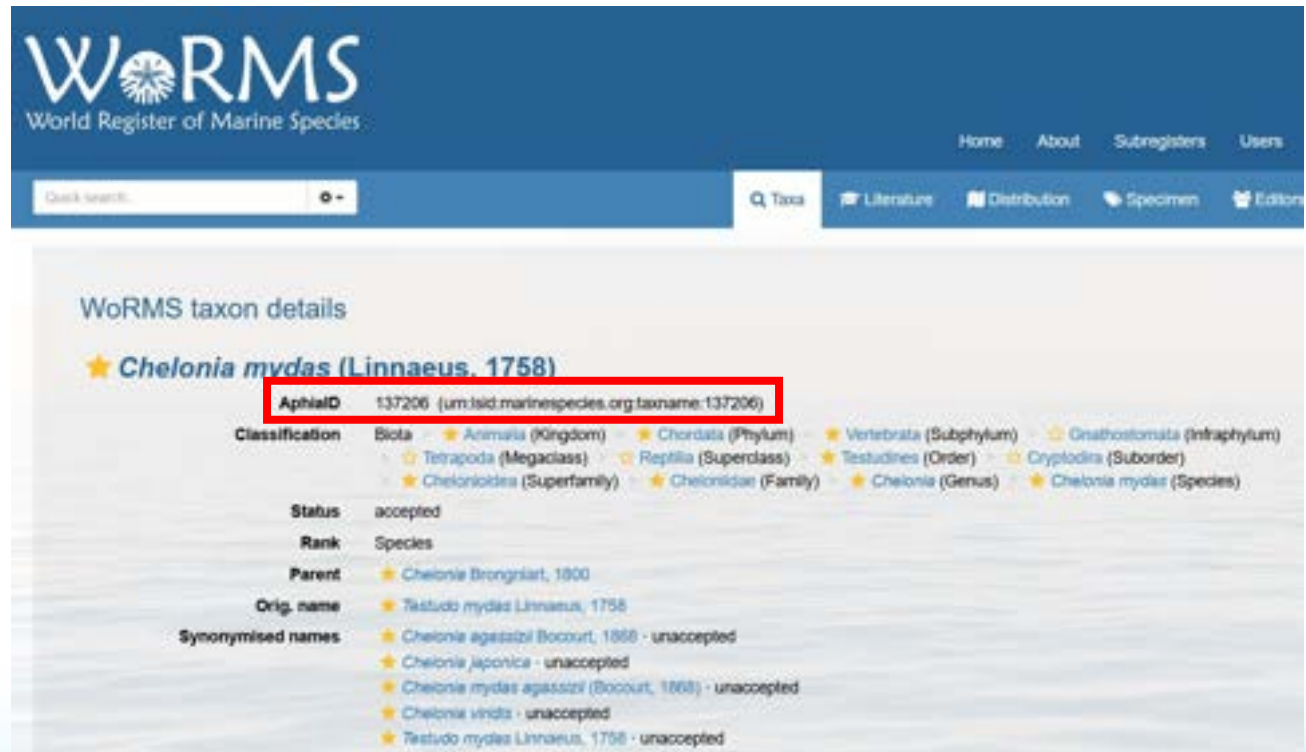
≡ Match names to an authoritative taxonomic register

≡ → Attach unique stable identifiers

≡ WoRMS LSIDs

≡ → Keep up with changing taxonomy

≡ → Avoid misspellings



The screenshot shows the WoRMS website interface. The header includes the WoRMS logo and navigation links: Home, About, Subregisters, Users. A search bar is present. The main content area displays 'WoRMS taxon details' for *Chelonia mydas* (Linnaeus, 1758). A red box highlights the 'AphiaID' field, which contains the value '137206' and the URL '(um:lsid.marinespecies.org:taxname:137206)'. Below this, a classification tree is shown, starting from Biota and branching through various taxonomic levels to the species level. The 'Status' is listed as 'accepted', and the 'Rank' is 'Species'. The 'Parent' is 'Chelonia Brongniart, 1800'. The 'Orig. name' is 'Testudo mydas Linnaeus, 1758'. A list of 'Synonymised names' is provided, including 'Chelonia agassizii Bocourt, 1866 - unaccepted', 'Chelonia japonica - unaccepted', 'Chelonia mydas agassizii (Bocourt, 1866) - unaccepted', 'Chelonia vinda - unaccepted', and 'Testudo mydas Linnaeus, 1758 - unaccepted'.

Field	Value
AphiaID	137206 (um:lsid.marinespecies.org:taxname:137206)
Classification	Biota → Animalia (Kingdom) → Chordata (Phylum) → Vertebrata (Subphylum) → Gnathostomata (Infraphylum) → Tetrapoda (Megaclass) → Reptilia (Superclass) → Testudines (Order) → Cryptodira (Suborder) → Chelonioidea (Superfamily) → Cheloniidae (Family) → Chelonia (Genus) → Chelonia mydas (Species)
Status	accepted
Rank	Species
Parent	Chelonia Brongniart, 1800
Orig. name	Testudo mydas Linnaeus, 1758
Synonymised names	Chelonia agassizii Bocourt, 1866 - unaccepted Chelonia japonica - unaccepted Chelonia mydas agassizii (Bocourt, 1866) - unaccepted Chelonia vinda - unaccepted Testudo mydas Linnaeus, 1758 - unaccepted

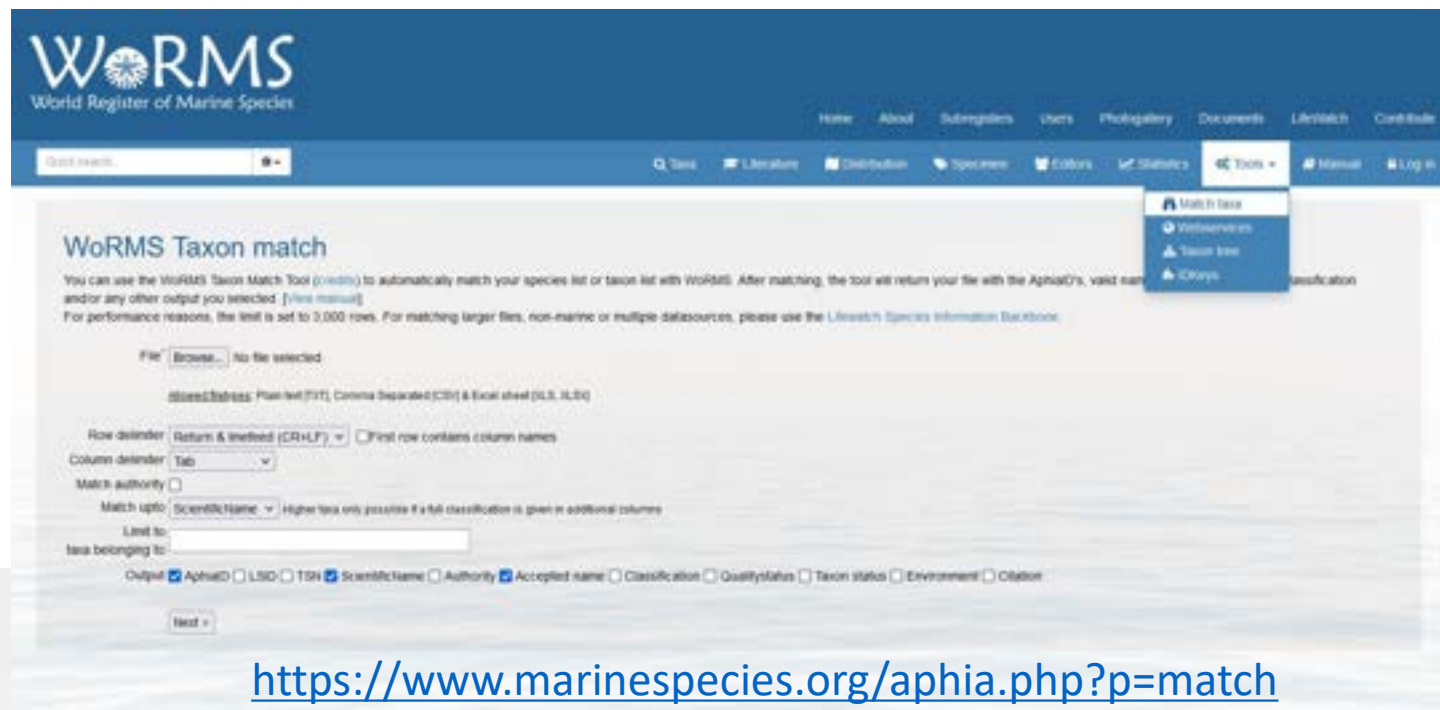
Taxonomic standardization

≡ WoRMS Taxon Match

≡ WoRMS Taxon Match Tool


≡ REST API

≡ Worrms (R client)



The screenshot shows the WoRMS Taxon Match web interface. The header includes the WoRMS logo and navigation links. The main content area is titled "WoRMS Taxon match" and provides instructions on how to use the tool. It includes a file upload section with a "Browse..." button and a "No file selected" message. Below this, there are options for "File format" (Plain text (TXT), Comma Separated (CSV), & Excel sheet (XLS, XLSX)), "Row delimiter" (Return & linefeed (CR+LF), First row contains column names), "Column delimiter" (Tab), "Match authority" (checkbox), "Match upto" (ScientificName, Higher taxa only possible if a full classification is given in additional columns), "Limit to" (text input), and " taxa belonging to" (text input). The "Output" section has checkboxes for AphiaID, LSO, TSN, ScientificName, Authority, Accepted name, Classification, QualityStatus, Taxon status, Environment, and Citation. A "Next" button is at the bottom.

<https://www.marinespecies.org/aphia.php?p=match>

WoRMS REST webservice 1.0.0 

<https://www.marinespecies.org/restapi-doc/swagger.json>

The definitions and operations are listed below. Click on an operation name to view it's details, and test it.

GET `/AphiaRecordsByMatchNames` Try to find AphiaRecords using the TAXAMATCH fuzzy matching algorithm by Tony Rees.

For each given scientific name (may include authority), try to find one or more AphiaRecords, using the TAXAMATCH fuzzy matching algorithm by Tony Rees. This allows you to (fuzzy) match multiple names in one call. Limited to 50 names at once for performance reasons.

<https://www.marinespecies.org/rest/>

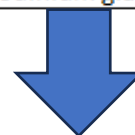
Taxonomic standardization

WoRMS Taxon Match

≡ Clean names before matching

- ≡ name for matching: only the scientific name of the taxon
- ≡ Other information should go in other fields

Name as provided	scientificName	scientificNameAuthor	lifeStage	sex	maximum length
Polysiphonia Greville, 1823	Polysiphonia	Greville, 1823			
Nephtys juv.	Nephtys		juv.		
Eupagurus pubescens zoea	Eupagurus pubescens		zoea		
Corbula crassa male adult	Corbula crassa		adult	male	
Katodinium glaucum <20um	Katodinium glaucum				20um



Use this column for the
taxon match

Taxonomic standardization

WoRMS Taxon Match

≡ Clean names before matching

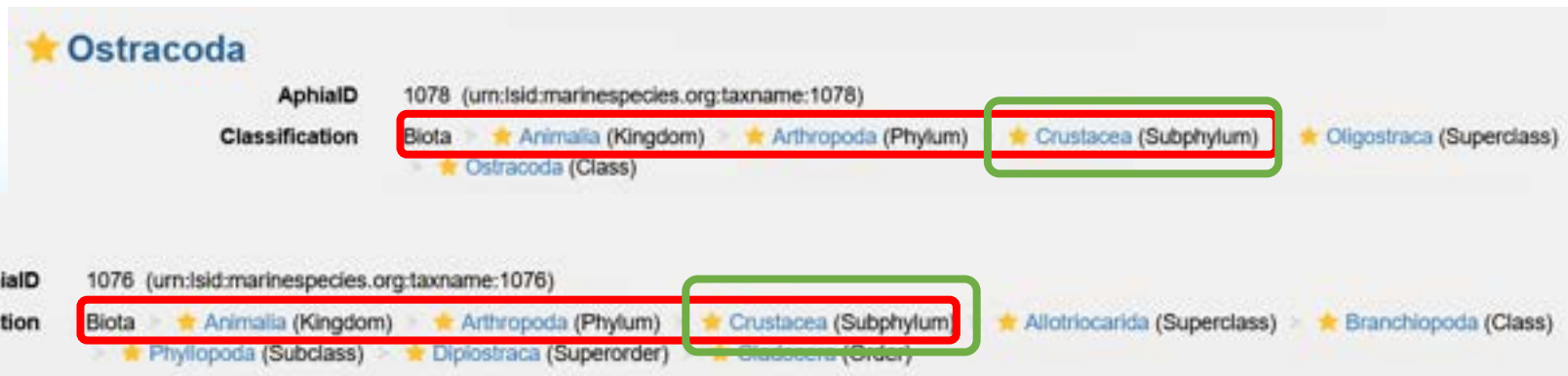
≡ scientificName: only the
scientific name of the taxon

≡ Other information should go in
other fields

≡ Uncertainty → scientificName:
the lowest taxonomic level at
which there is certainty

≡ *Cladocera/Ostracoda*
→ *Crustacea*

Name as provided	scientificName	scientificNameAuthor	lifeStage	sex	maximum length
Polysiphonia Greville, 1823	Polysiphonia	Greville, 1823			
Nephtys juv.	Nephtys		juv.		
Eupagurus pubescens zoea	Eupagurus pubescens		zoea		
Corbula crassa male adult	Corbula crassa		adult	male	
Katodinium glaucum <20um	Katodinium glaucum				20um



Taxonomic standardization

WoRMS Taxon Match

≡ Clean names before matching

- ≡ scientificName: only the scientific name of the taxon
- ≡ Other information should go in other fields
- ≡ Uncertainty → scientificName: the lowest taxonomic level at which there is certainty
 - ≡ *Cladocera/Ostracoda* → *Crustacea*
 - ≡ *Gadus cf. morhua* → *Gadus*
 - ≡ *Gadus morhua / macrocephalus* → *Gadus*
 - ≡ Mesozooplankton → *Animalia*

Name as provided	scientificName	scientificNameAuthor	lifeStage	sex	maximum length
Polysiphonia Greville, 1823	Polysiphonia	Greville, 1823			
Nephtys juv.	Nephtys		juv.		
Eupagurus pubescens zoea	Eupagurus pubescens		zoea		
Corbula crassa male adult	Corbula crassa		adult	male	
Katodinium glaucum <20um	Katodinium glaucum				20um

Taxonomic standardization

WoRMS Taxon Match

≡ Clean names before matching

- ≡ scientificName: only the scientific name of the taxon
- ≡ Other information should go in other fields
- ≡ Uncertainty → scientificName: the lowest taxonomic level at which there is certainty
 - ≡ identificationQualifier should contain the uncertain part (e.g. cf. *morhua*)
 - ≡ If it is not a taxonomic name, add it in taxonRemarks (e.g. mesozooplankton)

Name as provided	scientificName	scientificNameAuthor	lifeStage	sex	maximum length
Polysiphonia Greville, 1823	Polysiphonia	Greville, 1823			
Nephtys juv.	Nephtys		juv.		
Eupagurus pubescens zoea	Eupagurus pubescens		zoea		
Corbula crassa male adult	Corbula crassa		adult	male	
Katodinium glaucum <20um	Katodinium glaucum				20um

Name as provided	scientificName	identificationQualifier	taxonRemarks
Cladocera/Ostracoda	Crustacea	Cladocera/Ostracoda	
Gadus cfr. morhua	Gadus	cfr. morhua	
Mesozooplankton	Animalia		Mesozooplankton
Gadus morhua / macrocephalus	Gadus	morhua / macrocephalus	

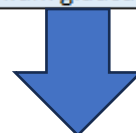
Taxonomic standardization

WoRMS Taxon Match

≡ Clean names before matching

- ≡ scientificName: only the scientific name of the taxon
- ≡ Other information should go in other fields
- ≡ Uncertainty → scientificName: the lowest taxonomic level at which there is certainty
- ≡ Name as provided can go in verbatimIdentification

Name as provided	scientificName	scientificNameAuthor	lifeStage	sex	maximum length
Polysiphonia Greville, 1823	Polysiphonia	Greville, 1823			
Nephtys juv.	Nephtys		juv.		
Eupagurus pubescens zoea	Eupagurus pubescens		zoea		
Corbula crassa male adult	Corbula crassa		adult	male	
Katodinium glaucum <20um	Katodinium glaucum				20um



verbatimIdentification



Name as provided	scientificName	identificationQualifier	taxonRemarks
Cladocera/Ostracoda	Crustacea	Cladocera/Ostracoda	
Gadus cfr. morhua	Gadus	cfr. morhua	
Mesozooplankton	Animalia		Mesozooplankton
Gadus morhua / macrocephalus	Gadus	morhua / macrocephalus	

Taxonomic standardization

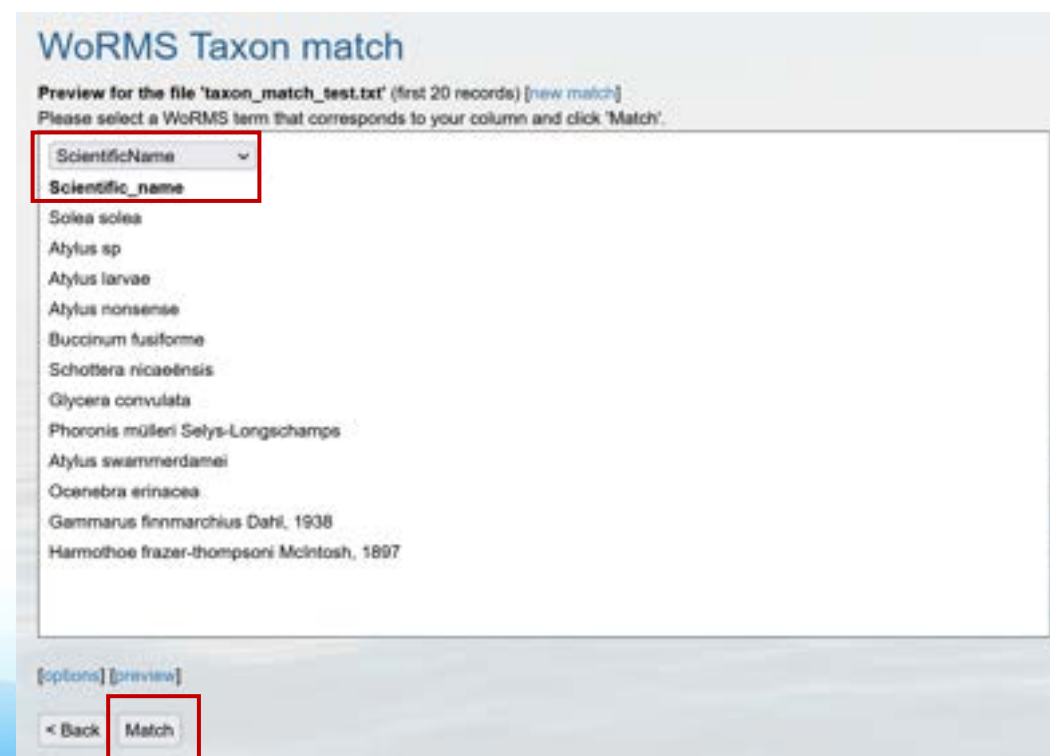
WoRMS Taxon Match

≡ WoRMS Taxon Match Tool

- ≡ Prepare file (Plain text [TXT], Comma Separated [CSV] & Excel Sheet [XLS, XLSX])
- ≡ For convenience => name the column with the cleaned name “Scientific_name” or “ScientificName”
- ≡ Upload onto website



```
File Edit Format View Help
Scientific_name
Solea solea
Atylus sp
Atylus larvae
Atylus nonsense
Buccinum fusiforme
Schottera nicaeensis
Glycera convulata
Phoronis mülleri Selys-Longschamps
Atylus swammerdamei
Ocenebra erinacea
Gammarus finnmarchius Dahl, 1938
Harmothoe frazer-thompsoni McIntosh, 1897
```



WoRMS Taxon match

Preview for the file 'taxon_match_test.txt' (first 20 records) [new match]
Please select a WoRMS term that corresponds to your column and click 'Match'.

ScientificName
Scientific_name

Solea solea
Atylus sp
Atylus larvae
Atylus nonsense
Buccinum fusiforme
Schottera nicaeensis
Glycera convulata
Phoronis mülleri Selys-Longschamps
Atylus swammerdamei
Ocenebra erinacea
Gammarus finnmarchius Dahl, 1938
Harmothoe frazer-thompsoni McIntosh, 1897

[options] [preview]

< Back Match

Taxonomic standardization

WoRMS Taxon Match

≡ Taxon match returns not only exact matches, also approximate matches

WoRMS Taxon match

Match preview for the file 'taxon_match_test.xlsx' - matching: 84.21% [[new match](#)]
If available, please select the [WoRMS](#) taxon that corresponds to your taxon. Then click 'Download'.

ScientificName	WoRMS match
Solea solea	Solea solea (Linnaeus, 1758)
Atylus sp	Atylus Leach, 1815
Atylus larvae	Atylus Leach, 1815
Atylus nonsense	(none)
Schottera nicaeensis	Schottera nicaeensis (J.V. Lamouroux ex Duby) Guiry & Hollenberg, 1975
Phoronis mülleri Selys-Longchamps	Phoronis muelleri Selys-Longchamps, 1903
Atylus swammerdami	Atylus swammerdami (H. Milne Edwards, 1830) accepted as Nototropis sv
Ocenebra erinacea	Ocenebra erinaceus (Linnaeus, 1758)
Gammarus finnmarchius Dahl, 1938	Gammarus finmarchicus Dahl, 1938 accepted as Echinogammarus incerta
Harmothoe frazer-thompsoni	Harmothoe fraserthompsoni McIntosh, 1897
Cuculus varius	Cuculus vicarius Röding, 1798 accepted as Conus locumtenens Blumenb.
Corbula crassa	Corbula crassa Reeve, 1843 accepted as Corbula ovalina Lamarck, 1818
Typhis montforti	Typhis montforti A. Adams, 1863 accepted as Monstrotypis montfortii (A.
Labidodemas leucopus	Labidodemas leucopus Haacke, 1880 accepted as Holothuria (Mertensioth
Holothuria (Mertensiothuria) hilla	Holothuria (Mertensiothuria) hilla Lesson, 1830

☐ Excel sheet (XLS)
 ☒ Excel sheet (XLSX)
 ☐ Text file
 ☐ SGML

[< Back](#)
[Download](#)

Taxonomic standardization

WoRMS Taxon Match

≡ Taxon match returns not only exact matches, also approximate matches

	A	B	C	D	E	F	G
1	ScientificName	AphiaID	Match type	LSID	TSN	Qualitystatus	Taxon status
2	Solea solea	127160	exact	urn:lsid:marinespecies.org:taxname:127160	173002	Checked by Taxonomic Editor	accepted
3	Atylus sp	101497	exact	urn:lsid:marinespecies.org:taxname:101497	93514	Checked by Taxonomic Editor	accepted
4	Atylus larvae	101497	exact	urn:lsid:marinespecies.org:taxname:101497	93514	Checked by Taxonomic Editor	accepted
5	Atylus nonsense						
6	Schottera nicae	494793	exact	urn:lsid:marinespecies.org:taxname:494793		Checked by Taxonomic Editor	unaccepted
7	Glycera convulata	155109	exact	urn:lsid:marinespecies.org:taxname:155109		Added by Database Management Team	unaccepted
8	Phoronis m	128549	phonetic	urn:lsid:marinespecies.org:taxname:128549	206663	Checked by Taxonomic Editor	accepted
9	Atylus swammerdamei	102131	phonetic	urn:lsid:marinespecies.org:taxname:102131	93523	Checked by Taxonomic Editor	accepted
10	Ocenebra erinacea	140405	near_1	urn:lsid:marinespecies.org:taxname:140405	73249	Checked by Taxonomic Editor	accepted
11	Gammarus finmarchius Dahl, 1938	102277	near_2	urn:lsid:marinespecies.org:taxname:102277	206449	Checked by Taxonomic Editor	accepted
12	Harmothoe frazer-thompsoni McIntosh, 1897	130764	near_2	urn:lsid:marinespecies.org:taxname:130764	64526	Checked by Taxonomic Editor	accepted

	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	ScientificName	Authority	AphiaID	ScientificName_accepted	Kingdom	Phylum	Class	Order	Family	Genus	Species	Citation	
2	Solea solea	(Linnaeus, 1758)	127160	Solea solea	Animalia	Chordata	Actinopterygii	Pleuronect	Soleidae	Solea	solea	Bailey, N. (2011). Sole	
3	Atylus	Leach, 1815	101497	Atylus	Animalia	Arthropoda	Malacostraca	Amphipod	Atylidae	Atylus		Lowry, J.; De Broyer,	
4	Atylus	Leach, 1815	101497	Atylus	Animalia	Arthropoda	Malacostraca	Amphipod	Atylidae	Atylus		Lowry, J.; De Broyer,	
5													
6	Schottera nicaeensis	(J.V.Lamouroux ex	145666	Schottera nicaeensis	Plantae	Rhodophyta	Florideophyceae	Gigartinales	Phyllophor	Schottera	nicaeensis	Guiry, M.D. (2011). Si	
7	Glycera convulata		130120	Glycera convulata	Animalia	Annelida	Polychaeta	Phyllodoce	Glyceridae	Glycera	convulata	WoRMS (2010). Glyc	
8	Phoronis muelleri	Selys-Lonchamps,	128549	Phoronis muelleri	Animalia	Phoronida				Phoronis	muelleri	Emig, C. (2011). Phor	
9	Atylus swammerdami	(Milne-Edwards, 18	102131	Atylus swammerdami	Animalia	Arthropoda	Malacostraca	Amphipod	Atylidae	Atylus	swammerdami	Costello, M.; Bellan-S	
10	Ocenebra erinaceus	(Linnaeus, 1758)	140405	Ocenebra erinaceus	Animalia	Mollusca	Gastropoda	Neogastrop	Muricidae	Ocenebra	erinaceus	Houart, R.; Gofas, S.	
11	Gammarus finmarchicus	Dahl, 1938	102277	Gammarus finmarchicus	Animalia	Arthropoda	Malacostraca	Amphipod	Gammarid	Gammarus	finmarchicus	Costello, M.; Bellan-S	
12	Harmothoe fraserthomsoni	McIntosh, 1897	130764	Harmothoe fraserthomsoni	Animalia	Annelida	Polychaeta	Phyllodoce	Polynoidae	Harmothoe	fraserthomsoni	Fauchald, K.; Barnich	

Taxonomic standardization

WoRMS Taxon Match

≡ Taxon match returns not only exact matches, also approximate matches

≡ WoRMS taxon match results:

- ≡ exact: all characters match exactly
- ≡ exact_subgenus: an exact match, but including the subgenus
- ≡ phonetic: sounds similar as, despite minor differences in spelling (soundex algorithm)
- ≡ near_1: perfect match, except for one character. This is a quite reliable match.
- ≡ near_2: good match, except for two characters. This needs an extra check.
- ≡ near_3: good match, except for three characters. This definitely needs an extra check.
- ≡ match_quarantine: match with a name that is currently in quarantine. Any name that has been used in the literature should in principle not be quarantined. So best to contact the WoRMS DMT about this.
- ≡ match_deleted: this is a match with a name that has been deleted and no alternative is available. Please contact the WoRMS DMT when you come across this.
- ≡ No match

≡ → Check and verify everything that is not an exact match...

Taxonomic standardization

WoRMS Taxon Match

≡ No match found

- ≡ Check if name was entered correctly
- ≡ Check if valid name → match with other registers:
 - ≡ LifeWatch taxon match
<https://www.lifewatch.be/data-services/>
- ≡ Check if the taxon is marine → lookup environment on IRMNG (<https://www.irmng.org/>)
 - ≡ Marine taxon: contact WoRMS DMT
 - ≡ Non-marine taxon:
 - ≡ Misidentification?
 - ≡ Not non-marine: contact WoRMS DMT

WoRMS Taxon match

Match preview for the file 'taxon_match_test.xlsx' - matching: 84.21% [new match]
If available, please select the [WoRMS](#) taxon that corresponds to your taxon. Then click 'Download'.

ScientificName	WoRMS match
Solea solea	Solea solea (Linnaeus, 1758)
Atylus sp	Atylus Leach, 1815
Atylus larvae	Atylus Leach, 1815
Atylus nonsense	(none)
Schottera nicaeensis	Schottera nicaeensis (J.V. Lamouroux ex Duby) Guiry & Hollenberg, 1975
Phoronis mülleri Selys-Longchamps	Phoronis muelleri Selys-Longchamps, 1903
Atylus swammerdamei	Atylus swammerdamei (H. Milne Edwards, 1830) accepted as Nototheniopsis
Ocenebra erinacea	Ocenebra erinacea (Linnaeus, 1758)
Gammarus finmarchicus Dahl, 1938	Gammarus finmarchicus Dahl, 1938 accepted as Echinogammarus incerta
Harmothoe frazer-thompsoni	Harmothoe frazerthompsoni McIntosh, 1897
Cuculus varius	Cuculus varius Röding, 1798 accepted as Conus locumtenens Blumenb.
Corbula crassa	Corbula crassa Reeve, 1843 accepted as Corbula ovalina Lamarck, 1816
Typhis montfortii	Typhis montfortii A. Adams, 1863 accepted as Monstrotaphis montfortii (A.
Labidodemas leucopus	Labidodemas leucopus Haacke, 1880 accepted as Holothuria (Mertensioth
Holothuria (Mertensiothuria) hilla	Holothuria (Mertensiothuria) hilla Lesson, 1830

☐ Excel sheet (XLS)
 ☒ Excel sheet (XLSX)
 ☐ Text file
 ☐ SGML

Taxonomic standardization

WoRMS Taxon Match

≡ Ambiguous matches (=multiple possible matches)

WoRMS Taxon match

Match preview for the file 'taxon_match_test.txt' - matching: 91.67% [[new match](#)]
If available, please select the WoRMS taxon that corresponds to your taxon. Then click 'Download'.

ScientificName	WoRMS match
Solea solea	Solea solea (Linnaeus, 1758)
Atylus sp	Atylus Leach, 1815
Atylus larvae	Atylus Leach, 1815
Atylus nonsense	(none)
Buccinum fusiforme	(ambiguous - select below)
Schottera nicaeensis	(ambiguous - select below)
Glycera convulata	Buccinum fusiforme Kiener, 1834 accepted as Buccinum humphreysianum Bennett, 1824 [exact]
Phoronis mülleri Selys-Longchamps	Buccinum fusiforme Broderip, 1830 accepted as Turrisipho fenestratus (W. Turton, 1834) [exact]
Atylus swammerdamei	Atylus swammerdamei (H. Milne Edwards, 1830) accepted as Nototropis swammerdamei
Ocenebra erinacea	Ocenebra erinacea (Linnaeus, 1758)
Gammarus finmarchius Dahl, 1938	Gammarus finmarchius Dahl, 1938 accepted as Echinogammarus incertae sedis finma
Harmothoe frazer-thompsoni McIntosh, 1897	Harmothoe frazerthompsoni McIntosh, 1897

< >

☐ Excel sheet (XLS) ☐ Excel sheet (XLSX) ☒ Text file ☐ SGML

< Back Download

Taxonomic standardization

WoRMS Taxon Match

≡ Ambiguous matches (=multiple possible matches)

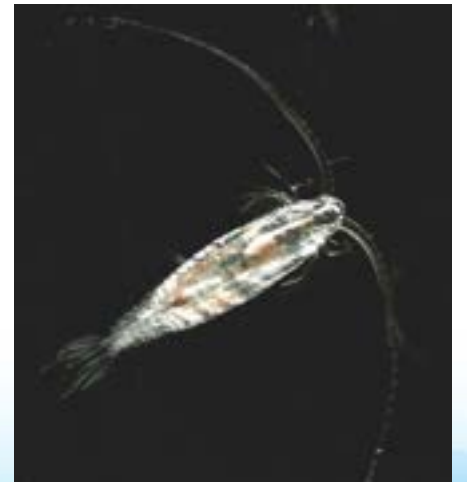
≡ Check authority

≡ Check classification

Chondracanthus Kützing, 1843
Kingdom Plantae (Rhodophyta)



Chondracanthus Delaroche, 1811
Kingdom Animalia (Crustacea)



Taxonomic standardization

WoRMS Taxon Match

≡ Ambiguous matches (=multiple possible matches)

≡ Check authority

≡ Check classification

≡ After resolving ambiguous matches

→ download results

WoRMS Taxon match

Match preview for the file 'taxon_match_test.txt' - matching: 91.67% [[new match](#)]
If available, please select the [WoRMS](#) taxon that corresponds to [your](#) taxon. Then click 'Download'.

ScientificName	WoRMS match
Solea solea	Solea solea (Linnaeus, 1758)
Atylus sp	Atylus Leach, 1815
Atylus larvae	Atylus Leach, 1815
Atylus nonsense	(none)
Buccinum fusiforme	(ambiguous - select below)
Schottera nicaeensis	(ambiguous - select below)
Glycera convulata	Buccinum fusiforme Kiener, 1834 accepted as Buccinum humphreysianum Bennett, 1824 [exact]
Phoronis mülleri Selys-Longchamps	Buccinum fusiforme Broderip, 1830 accepted as Turrisipho fenestratus (W. Turton, 1834) [exact]
Atylus swammerdamei	Atylus swammerdamei (H. Milne Edwards, 1830) accepted as Nototropis swammerdamei
Ocenebra erinacea	Ocenebra erinacea (Linnaeus, 1758)
Gammarus finmarchius Dahl, 1938	Gammarus finmarchius Dahl, 1938 accepted as Echinogammarus incoertae sedis finma
Harmothoe frazer-thompsoni McIntosh, 1897	Harmothoe frazerthompsoni McIntosh, 1897

< >

☐ Excel sheet (XLS) ☐ Excel sheet (XLSX) ☒ Text file ☐ SGML

< Back Download

Taxonomic standardization

WoRMS Taxon Match

	A	B	C	D	E	F	G
1	ScientificName	AphiaID	Match type	LSID	TSN	Qualitystatus	Taxon status
2	Solea solea	127160	exact	urn:lsid:marinespecies.org:taxname:127160	173002	Checked by Taxonomic Editor	accepted
3	Atylus sp	101497	exact	urn:lsid:marinespecies.org:taxname:101497	93514	Checked by Taxonomic Editor	accepted
4	Atylus larvae	101497	exact	urn:lsid:marinespecies.org:taxname:101497	93514	Checked by Taxonomic Editor	accepted
5	Atylus nonsense						
6	Schottera nicae	494793	exact	urn:lsid:marinespecies.org:taxname:494793		Checked by Taxonomic Editor	unaccepted
7	Glycera convulata	155109	exact	urn:lsid:marinespecies.org:taxname:155109		Added by Database Management Team	unaccepted
8	Phoronis m	128549	phonetic	urn:lsid:marinespecies.org:taxname:128549	206663	Checked by Taxonomic Editor	accepted
9	Atylus swammerdamei	102131	phonetic	urn:lsid:marinespecies.org:taxname:102131	93523	Checked by Taxonomic Editor	accepted
10	Ocenebra erinacea	140405	near_1	urn:lsid:marinespecies.org:taxname:140405	73249	Checked by Taxonomic Editor	accepted
11	Gammarus finnmarchius Dahl, 1938	102277	near_2	urn:lsid:marinespecies.org:taxname:102277	206449	Checked by Taxonomic Editor	accepted
12	Harmothoe frazer-thompsoni McIntosh, 1897	130764	near_2	urn:lsid:marinespecies.org:taxname:130764	64526	Checked by Taxonomic Editor	accepted

≡ **LSID** → DwC field **scientificNameID**

Value standardization

- ≡ How to standardize depends on the field
- ≡ Let's look into more detail into:
 - ≡ Taxonomy
 - ≡ **Geography**
 - ≡ Time

Geographic standardization

Coordinates

Different spatial reference systems

OBIS:

Decimal degrees

EPSG:4326 (WGS84)

geodeticDatum	
Identifier	http://rs.tdwg.org/dwc/terms/geodeticDatum
Definition	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in dwcdecimalLatitude and dwcdecimalLongitude are based.
Comments	Recommended best practice is to use the EPSG code of the 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 <i>unknown</i> . 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.
Examples	<div>EPSG:4326</div> <div>WGS84</div> <div>NAD27</div> <div>Campo Inchauspe</div> <div>European 1950</div> <div>Clarke 1866</div> <div>unknown</div>

Geographic standardization

Coordinates

Names




Marineregions.org
towards a standard for geo-referenced marine names

About Gazetteer Maritime Boundaries Sources Statistics Downloads

Marine Gazetteer Search results
Your search for 'north sea' returned 25 matching records

Search
Browse
About
Tutorial
Webservices
Login

- [Belgian part of the North Sea](#) (Marine Region) has preferred alternative [Belgian Exclusive Economic Zone](#)
- [Central North Sea](#) (General Sea Area)
- [Danish part of the North Sea](#) (Marine Region)
- [Dutch part of the North Sea](#) (Marine Region) has preferred alternative [Dutch Exclusive Economic Zone](#)
- [EMODnet Biology Reporting Area: North Sea](#) (EMODnet Biology Reporting Area)
- [French part of Greater North Sea](#) (MSFD Marine subregions)
- [French part of the north sea](#) (Marine Region)
- [German part of the North Sea](#) (Marine Region)
- [Greater North Sea](#) (ICES Ecoregion)
- [Greater North Sea, incl. the Kattegat and the English Channel](#) (MSFD Marine subregions)
- [North Sea](#) (IHO Sea Area)
- [North Sea](#) (Large Marine Ecosystem) has preferred alternative [North Sea](#)
- [North Sea](#) (Marine Ecoregion of the World (MEOW)) has preferred alternative [North Sea](#)
- [North Sea](#) (ICES Ecoregion) replaced by [Greater North Sea](#)
- [North Sea](#) (SeaVol SeaArea - level 3) has preferred alternative [North Sea](#)
- [North Sea](#) (SeaVol SeaArea - sub-region) has preferred alternative [North Sea](#)
- [North Sea Bottom Current](#) (Current)
- [North Sea Canal](#) (Channel)
- [North Sea Channel](#) (Seachannel)
- [Northern North Sea](#) (General Sea Area)
- [Northern Part of the North Sea](#) (General Sea Area)
- [Norwegian part of the North Sea](#) (Marine Region)
- [Southern Right of the North Sea](#) (Right)
- [Southern North Sea](#) (General Sea Area)
- [United Kingdom part of the North Sea](#) (Marine Region)



Marineregions.org
towards a standard for geo-referenced marine names

About Gazetteer Maritime Boundaries Sources Statistics Downloads

Marine Gazetteer Placedetails

Search
Browse
About
Tutorial
Webservices

MRID: <http://marineregions.org/mrid/2300>

Status: Proposed standard

Names	Language	Name	Name source
English	North Sea	(1953). Limits of oceans and seas. 3rd edition. IHO Special Publication, 23. International Hydrographic Organization (IHO). Monaco. 38 pp. (look up in BIB)	
Dutch	Noordzee		

PlaceType: IHO Sea Area

Latitude: 56° 25' 26.4" N (56.4239952°)

Longitude: 2° 44' 16.3" E (2.73796024°)

Precision: 711089 meter

Min. Lat: 50° 58' 43.3" N (50.9954°)

Min. Long: 4° 20' 43.3" W (-4.4454°)


Max. Lat: 61° 1' 1.3" N (61.017°)

Max. Long: 12° 0' 21.4" E (12.0059°)

Source: (1953). Limits of oceans and seas. 3rd edition. IHO Special Publication, 23. International Hydrographic Organization (IHO). Monaco. 38 pp. (look up in [BIB](#))

Relation: Part of [North Atlantic Ocean](#) (IHO Sea Area) [View hierarchy](#)

Map



Download: Layer: [MarineRegions.shp](#) - format: [GML3](#) [Download](#)

Shapefile: [download](#) or view the complete IHO Sea Area shapefile

Edit history: Last edited on 2017-01-18 17:22:03 by [Da Happoni](#) [View history](#)

(Source: (Source: edited) (Source: merged)

Geographic standardization

≡ Coordinates = basis of a biogeographic information system

≡ When no coordinates are provided...

≡ Derive from other location information

Geographic standardization

≡ OBIS Maptool:

<https://obis.org/maptool/>

≡ get latitude, longitude and radius for a geographic area (polygon) or a transect (line) drawn on map

Layers

Switch layers on or off. Layers from Marine Regions.

EEZ boundaries

WKT

Generate WKT

WKT

Coordinates

Add a location using decimal longitude and latitude (space or comma separated).

Enter coordinates Add

Geocoding

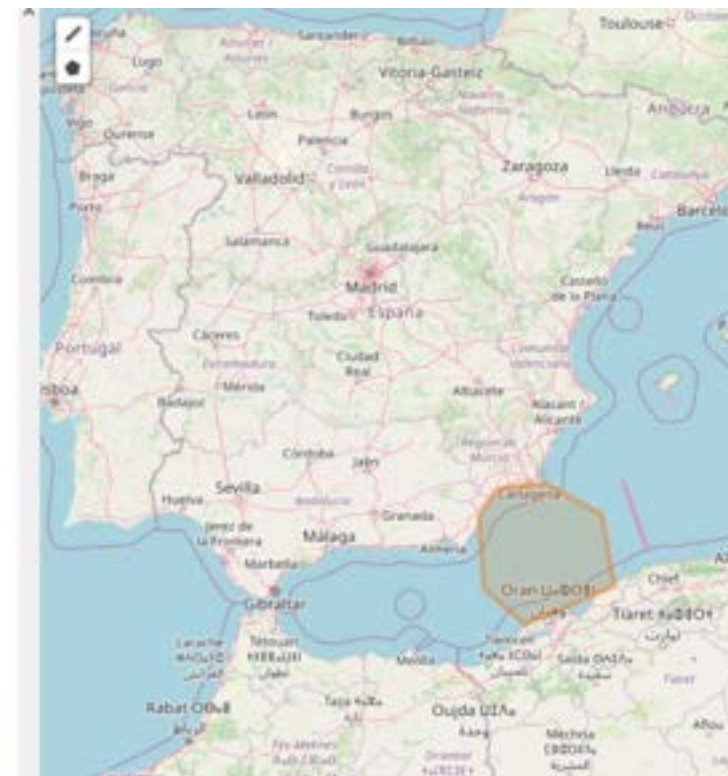
Find locations by name and add them to the locations list. Geocoding by Marine Regions.

Enter location name Submit Clear

Type	Name	Longitude	Latitude	Uncertainty (m)
No results				

Locations

	Longitude	Latitude	Radius (m)	Name	Shore distance (m)	Depth (m)	
<input type="checkbox"/>	-0.8752	36.7641	136,200				✕
<input type="checkbox"/>	-0.8361	37.1767	51,725				✕



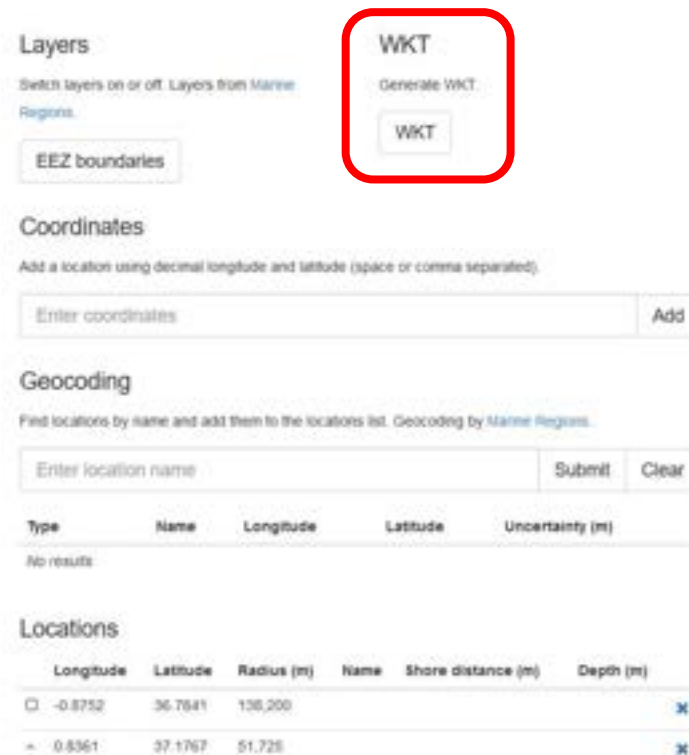
Geographic standardization

≡ OBIS Maptool:

<https://obis.org/maptool/>

≡ get latitude, longitude and radius **for a geographic area (polygon) or a transect (line) drawn on map**

≡ Shape → footprintWKT



Layers
Switch layers on or off. Layers from Marine Regions.
EEZ boundaries

WKT
Generate WKT
WKT

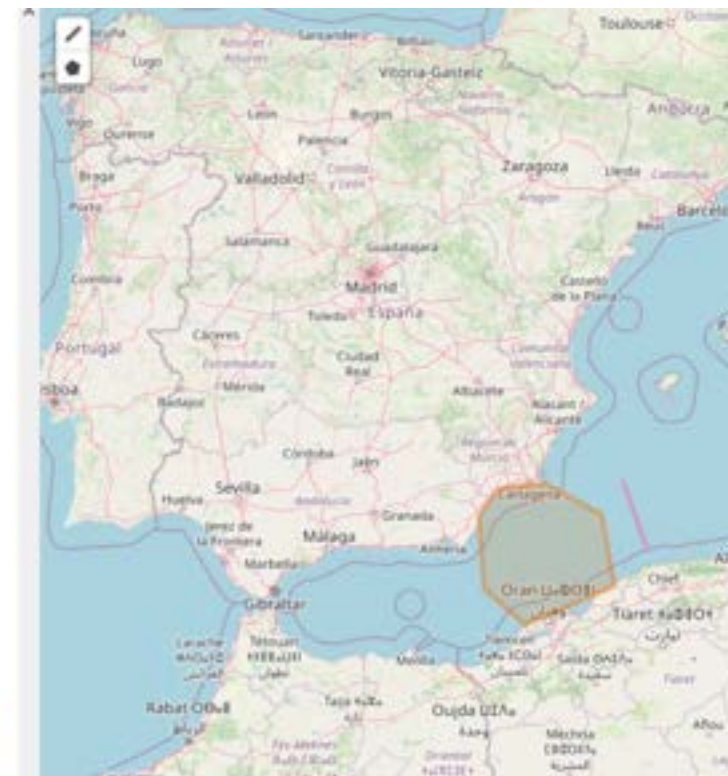
Coordinates
Add a location using decimal longitude and latitude (space or comma separated).
Enter coordinates Add

Geocoding
Find locations by name and add them to the locations list. Geocoding by Marine Regions.
Enter location name Submit Clear

Type	Name	Longitude	Latitude	Uncertainty (m)
No results				

Locations

	Longitude	Latitude	Radius (m)	Name	Shore distance (m)	Depth (m)
□	-0.8752	36.7641	136,200			✕
△	-0.8361	37.1767	51,725			✕



Geographic standardization

≡ OBIS Maptool:

<https://obis.org/maptool/>

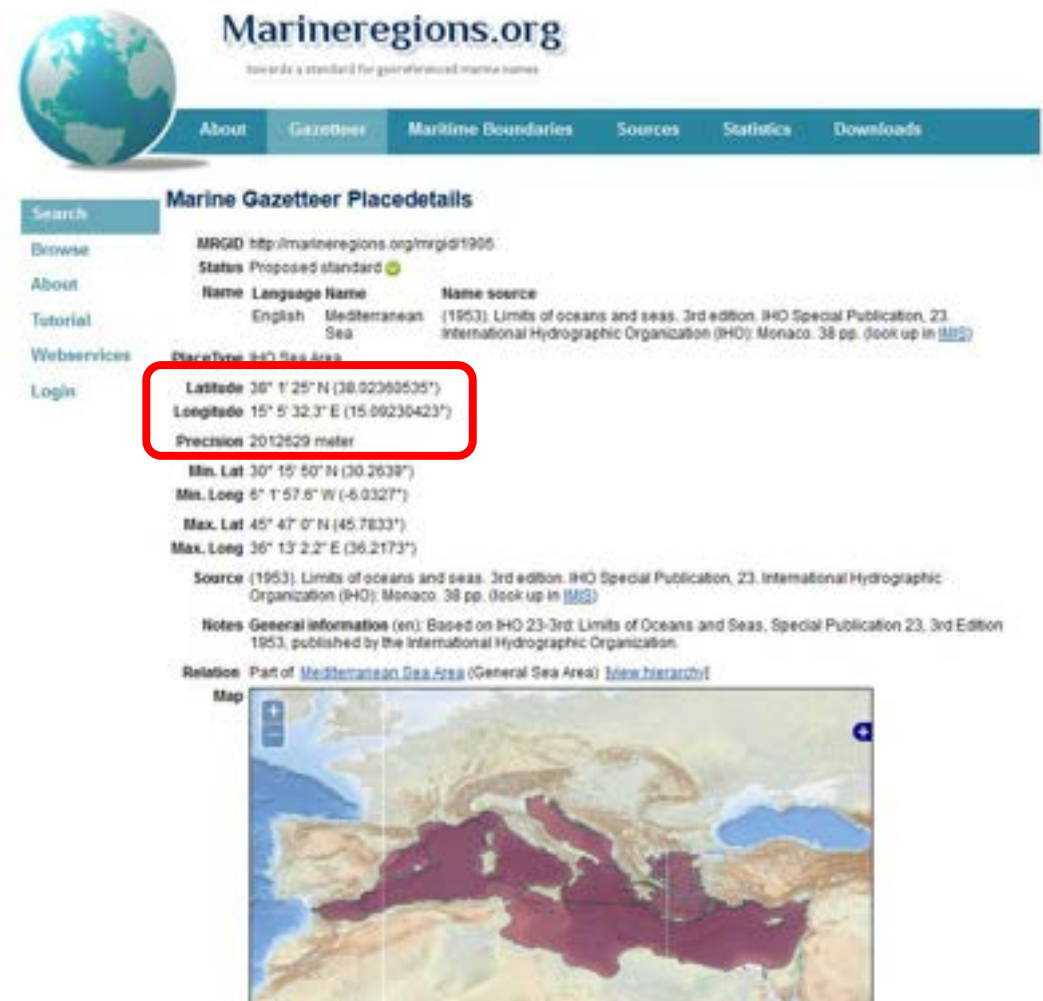
≡ get latitude, longitude and radius **for a geographic area (polygon) or a transect (line) drawn on map**

≡ Shape → footprintWKT

≡ Marine Regions Gazetteer:

<https://marineregions.org/>

≡ get latitude, longitude and precision **based on a place name**



Marineregions.org
towards a standard for geographical marine names

About Gazetteer Maritime Boundaries Sources Statistics Downloads

Search

Marine Gazetteer Placedetails

MRGID: <http://marineregions.org/mrgid/1905>

Status: Proposed standard

Name	Language	Name	Name source
English	Mediterranean Sea	(1953). Limits of oceans and seas. 3rd edition. IHO Special Publication, 23. International Hydrographic Organization (IHO): Monaco. 38 pp. (look up in MIS)	

Definition: 34°N Sea Area

Latitude: 36° 1' 25" N (38.02360535°)

Longitude: 15° 5' 32.3" E (15.09236423°)

Precision: 2012629 meter

Min. Lat: 30° 15' 50" N (30.2639°)

Min. Long: 6° 1' 57.6" W (-6.0327°)

Max. Lat: 45° 47' 0" N (45.7833°)

Max. Long: 36° 13' 2.2" E (36.2173°)

Source: (1953). Limits of oceans and seas. 3rd edition. IHO Special Publication, 23. International Hydrographic Organization (IHO): Monaco. 38 pp. (look up in [MIS](#))

Notes: General information (en): Based on IHO 23-3rd. Limits of Oceans and Seas, Special Publication 23, 3rd Edition 1953, published by the International Hydrographic Organization.

Relation: Part of [Mediterranean Sea Area](#) (General Sea Area) [View hierarchy](#)

Map

Geographic standardization

≡ OBIS Maptool:

<https://obis.org/maptool/>

≡ get latitude, longitude and **radius** for a geographic area (polygon) or a transect (line) drawn on map

≡ Shape → footprintWKT

≡ Marine Regions Gazetteer:

<https://marineregions.org/>

≡ get latitude, longitude and **precision** based on a place name

≡ Do not refer to 'uncertain' locations as points, but as areas

≡ → Include **coordinateUncertaintyInMeters** !

Geographic standardization

≡ When no coordinates are provided...

- ≡ Derive from other location information
 - ≡ But take care! ""To georeference poorly is worse than not to georeference at all."
 - ≡ Georeferencing Best Practices:
<https://docs.gbif.org/georeferencing-best-practices/1.0/en/>
 - ≡ Georeferencing Quick Reference Guide:
<https://docs.gbif.org/georeferencing-quick-reference-guide/1.0/en/>

≡ Do not refer to 'uncertain' locations as points, but as areas

≡ → Include
coordinateUncertaintyInMeters
!

Value standardization

- ≡ How to standardize depends on the field
- ≡ Let's look into more detail into:
 - ≡ Taxonomy
 - ≡ Geography
 - ≡ **Time**

Temporal standardization

- ≡ Standard: ISO 8601-1:2019
- ≡ YYYY-MM-DD
- ≡ No timezone specified → local time
 - ≡ If UTC: add a Z at the end
- ≡ Unknown time → do not add time (do not use 00:00)

eventDate	
Identifier	http://rs.tdwg.org/dwc/terms/eventDate
Definition	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.
Comments	Recommended best practice is to use a date that conforms to ISO 8601-1:2019.

≡ Examples:

≡ Dates:

- ≡ 1948-09-13
- ≡ 1993-01/02
- ≡ 1993-01
- ≡ 1993

≡ Dates with Specific Times:

- ≡ 1973-02-28T15:25:00
- ≡ 2008-04-25T09:53

≡ Dates with Time Zones:

- ≡ 2005-08-31T12:11+12
- ≡ 2013-02-16T04:28Z

≡ Date and Time Intervals:

- ≡ 1993-01-26T04:39+12/1993-01-26T05:48+12

Value standardization

- ≡ How to standardize depends on the field
- ≡ Let's look into more detail into:
 - ≡ Taxonomy
 - ≡ Geography
 - ≡ Time
- ≡ Measurements

lifeStage

Identifier	http://rs.tdwg.org/dwc/iri/lifeStage
Definition	The age class or life stage of the dwc:Organism(s) at the time the dwc:Occurrence was recorded.
Comments	Recommended best practice is to use a controlled vocabulary. Terms in the dwciri namespace are intended to be used in RDF with non-literal objects.

individualCount

Identifier	http://rs.tdwg.org/dwc/terms/individualCount
Definition	The number of individuals present at the time of the dwc:Occurrence.

Measurements standardization

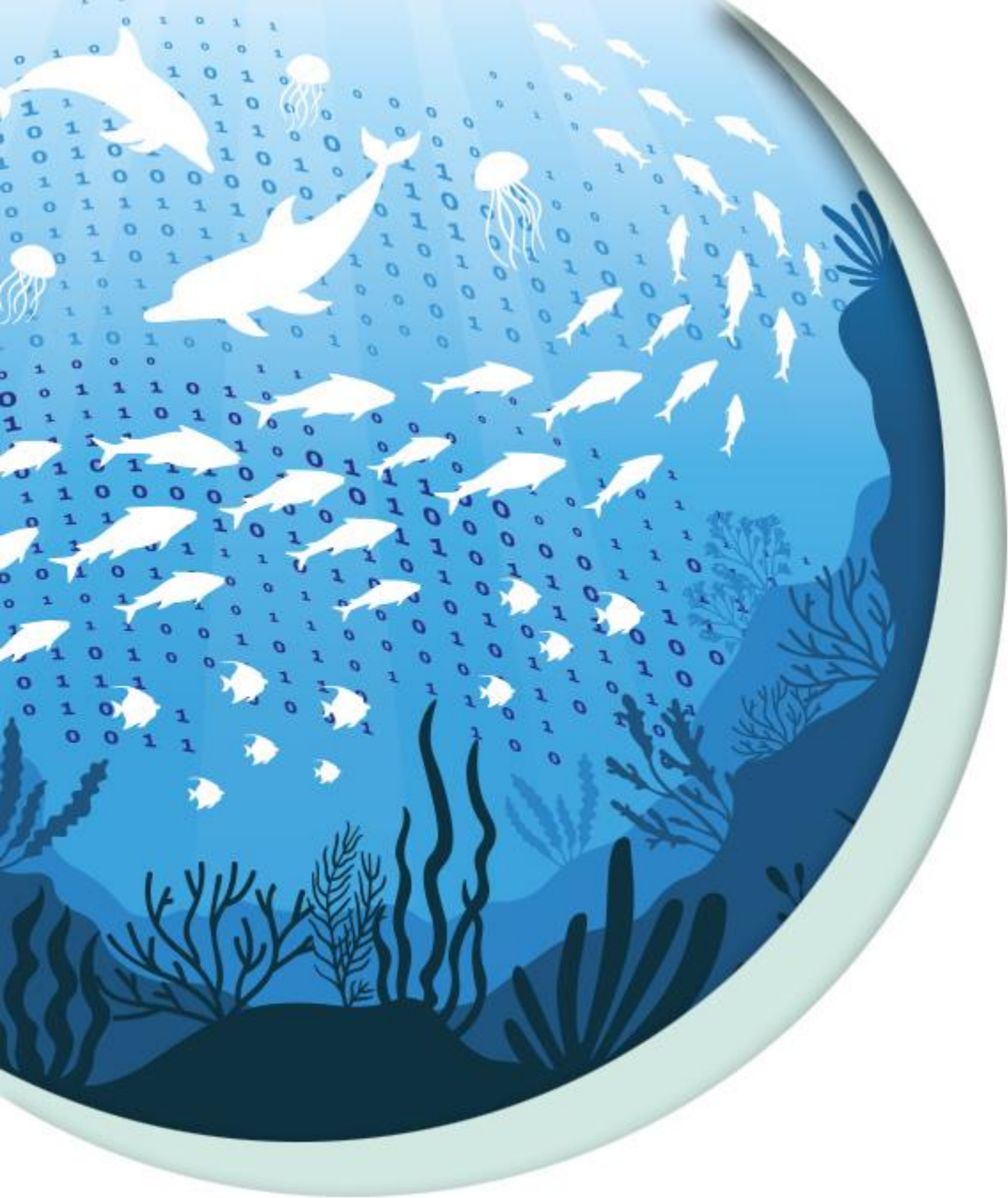
- ≡ Not all measurement types have a corresponding DwC term
- ≡ Solution → eMOF (extendedMeasurementOrFact)
- ≡ BODC NERC vocabulary
 - ≡ Standardise parameter names, units and values
- ≡ Details in the eMOF presentation

	A	B	C	D	E	F	G	H
1	EventID	Occurrence measurementType	measurementTypeID	measurementValue	measurementValueID	measurementUnit	measurementUnitID	
2	BIOFUN1_CSIC_BIOI	Abundance of biological entity specified elsewhere per unit area of the bed	http://vocab.nerc.ac.uk/collection/P01/current/SDBIOL02/	0.00025		N/km2	http://vocab.nerc.ac.uk/collection/P06/current/NPKM/	
3	BIOFUN1_CSIC_BIOI	Count (in assayed sample) of biological entity specified elsewhere	http://vocab.nerc.ac.uk/collection/P01/current/OCOUNT01/	1				
4	BIOFUN1_CSIC_BIOI	Wet weight biomass of biological entity specified elsewhere per unit area of the bed	http://vocab.nerc.ac.uk/collection/P01/current/SDBIOL05/	0.00375		kg/km2	http://vocab.nerc.ac.uk/collection/P06/current/KGAK/	
5	BIOFUN1_CSIC_BIOI	Abundance of biological entity specified elsewhere per unit area of the bed	http://vocab.nerc.ac.uk/collection/P01/current/SDBIOL02/	0.00025		N/km2	http://vocab.nerc.ac.uk/collection/P06/current/NPKM/	
6	BIOFUN1_CSIC_BIOI	Count (in assayed sample) of biological entity specified elsewhere	http://vocab.nerc.ac.uk/collection/P01/current/OCOUNT01/	1				
7	BIOFUN1_CSIC_BIOI	Wet weight biomass of biological entity specified elsewhere per unit area of the bed	http://vocab.nerc.ac.uk/collection/P01/current/SDBIOL05/	0.00675		kg/km2	http://vocab.nerc.ac.uk/collection/P06/current/KGAK/	
8	BIOFUN1_BF1A01	Sampling device aperture length	http://vocab.nerc.ac.uk/collection/Q01/current/Q0100014/	2.5		m	http://vocab.nerc.ac.uk/collection/P06/current/ULAA/	
9	BIOFUN1_BF1A01	Sampling device aperture width	http://vocab.nerc.ac.uk/collection/Q01/current/Q0100013/	1.2		m	http://vocab.nerc.ac.uk/collection/P06/current/ULAA/	
10	BIOFUN1_BF1A01	Sampling instrument name	http://vocab.nerc.ac.uk/collection/Q01/current/Q0100002/	Agassiz dredge		http://vocab.nerc.ac.uk/collection/L22/current/TOOL1252/		
11	BIOFUN1_BF1A01	Sampling net mesh size	http://vocab.nerc.ac.uk/collection/Q01/current/Q0100015/	12		mm	http://vocab.nerc.ac.uk/collection/P06/current/UXMM/	
12	BIOFUN1_BF1A01	Speed of measurement platform relative to ground surface (speed over ground)	http://vocab.nerc.ac.uk/collection/P01/current/APSAZ01/	2		knots	http://vocab.nerc.ac.uk/collection/P06/current/UKNT/	

Example eMOF table (example from OTGA course “Contributing datasets to EMODnet Biology”)

Resources

- ≡ <https://dwc.tdwg.org/terms/>
- ≡ https://manual.obis.org/common_formatissues.html#spatial
- ≡ <https://docs.gbif.org/georeferencing-best-practices/1.0/en/>
- ≡ <https://docs.gbif.org/georeferencing-quick-reference-guide/1.0/en/>
- ≡ https://www.marinespecies.org/tutorial_taxonmatch.php
- ≡ https://manual.obis.org/common_formatissues.html#temporal-dates-and-times



DTO-BioFlow

Integration of biodiversity monitoring
data into the Digital Twin Ocean

THANKS!