Applying Linked Data Principles to Static ISO Documents

Adam Shepherd, Technical Director BCO-DMO



Biological & Chemical Oceanography Data Management Office

2019 ESIP Winter Meeting January 16, 2019



BC - DM

Biological & Chemical Oceanography Data Management Office





How do we enable static metadata records to evolve over time with new knowledge?

Example: Measuring "Photosynthetic Rate"

Dataset Name	Brief Description	Project	PI-Supplied Parameter Name
Host-symbiont reinfections	Symbiont removed and anemone reinfected with one of two strains of Symbiodinium	AnemoneOA	photo_output_symb
Imaging Pulse Amplitude Modulator Fluorometer Data	Imaging Pulse Amplitude Modulator Fluorometer Data	Resilient Acerv	mYield
Photosynthesic parameters - Rosario and Varadero	Photosynthesic parameters - Rosario and Varadero 2016 and 2017	Varadero Reef	alpha, Pmax_n, Pmax_g
Photosynthesis- Irradiance curve (P-E)	Photosynthesis-Irradiance curve	Varadero Reef	Slope_O2_prod, Rate_O2_prod
Photosynthetic and calcification rates	Photosynthetic and calcification rates of Pleurochrysis carterae	OA_Copes_Coccoliths	photo_rate, photoRate_LDcycle, photoRate_cellDensity_ugC, photoRate_cellDensity_pmoIC
respiration and photosynthesis II	Biomass-normalized dark respiration and net photosynthesis rates in coral and algae as a function of pCO2 where LEDR (light-enhanced dark respiration) was measured	OA coral adaptation	photosynthesis
Thalassiosira pseudonana cyclostat	Thalassiosira pseudonana cyclostat experimental results	Stressors on Marine Phytoplankton	PI

Definition added in May 2016

• Found in 7 Datasets

- 1st Dataset published May 2016
 - Definition changed June 2016

https://www.bco-dmo.org/parameter/648663

Updating "Photosynthetic Rate"

Photosynthesis is a process by which plants and other organisms use light energy and chlorophyll to convert water, carbon dioxide, and minerals into oxygen and energy-rich organic compounds. Photosynthesis is one of the mechanisms responsible for Primary Production. The methods used to estimate rate of photosynthesis vary widely as do the units of measurement, although estimates of photosynthesis are often reported as micromol CO2 m-2 s-1 or some variation of that

I took out this part: "although estimates of photosynthesis are often reported as micromol CO2 m-2 s-1 or some variation of that". CO2 is not a product of photosynthesis so when you measure photosynthesis you are not measuring CO2. You are measuring C or O2 usually. - mda, 6/13/2016

A compilation of dissolved noble gas and N2/Ar ratio measurements collected from 1999-2016 in locations spanning the globe

	Observations and Measurements	
	Fait Names List	
Passade	Bearigners	-
100,010	a pegantel screenal literiteris-lach colas, crigar is incluidase	-
11.00_1000	a alling according of the CDFC number or latest feet. Recording scatters, followed by collegate online names, followed by a setter, followed by for aligneeting	
tonio .	Realized Tealer in Space Last.	
angenes.	To implate it To depart regress Last, Repaire nation will do ingress line:	-
-	canine of its water that he water perights constrained free. Street a cash when water cash is constrained cash origin for the	
-	cardiar of the spitch that for white carging even three rises. We can avoid white card white the spit- te on card an income white cardiar fails white cards constraint at a given with reconstraint.	
	namber efficienzat al an inderstaat onder has the excita templometer dater fact, foal is used where he fight make excernal are grow dater writer.	-
	satise of to stary lighter hadings dier for to adde satisfies and light for.	-
pine .	per againe fragme, north, and day observe are he reacted damp in semicirceller.	-
	no.et, signifier they are needly, and day, and work the Bendlet that surrights were called and	
-	ing legal a house, with and legislamican he cannot any a measurable.	-
	pressore that	-
(Tites)	and branches respect to be 210 relation 2	-

View/Open

data_global-noble-gases.tsv (1.120Mb)
ISO19115-2.xml (338.3Kb)
Field_names.pdf (29.75Kb)
Global_Hammeetal2019.mat (400.0Kb)
Readme_Hammeetal2019.txt (20.55Kb)
Dataset_description.pdf (197.0Kb)

Date

2018-08-24

Author

Hamme, Roberta C. ⊘ Jenkins, William J. ⊘

Citable URI

https://hdl.handle.net/1912/10538

Date Created

2018-08-13

Location

North Atlantic Eastern Tropical Pacific - Transect from Peru to Tahiti Oxygen minium zone; East Pacific Rise Global oceans

westlimit: -159.9952; southlimit: -68.1081; eastlimit: 178.9985; northlimit: 78.9988

DOI

10.1575/1912/bco-dmo.744563

Inert gases dissolved in the ocean are powerful tracers of the impact of physical processes on gases, particularly air-sea gas exchange (by both diffusive and bubblemeditated processes), temperature change, atmospheric pressure variation, mixing between different water masses, and ice processes. We have compiled a global ocean database of dissolved neon, argon, and krypton measurements, supplemented by helium, xenon, and nitrogen/argon (N2/Ar) ratios in some locations. Samples were collected on board multiple research cruises spanning the period 1999 through 2016 and analyzed by mass spectrometry at four different shore-based laboratories (University of Victoria, Woods Hole Oceanographic Institution, University of Washington, and Scripps Institution of Oceanography). For a complete list of measurements, refer to the supplemental document 'Field_names.pdf', and a full dataset description is included in A compilation of dissolved noble gas and N2/Ar ratio measurements collected from 1999-2016 in locations spanning the globe

	Observations and Measurements			
Fall Names List				
Passade	Becquin	-		
100,010	a pegantel screenal literiteris-lach colas, crigar is incluidase	-		
11100,1000	s alling according of the CDPC number is baseline. Bitmatility scatter, followed by collegate online names, followed by a solar, followed by Tradigments			
ten or	Realized Tealer in Space Last.	-		
traphone .	To implate it To description (and Repairs) rations with the improve lites:	-		
	canine of its water that he water perights constrained free. Street a cash when water cash is constrained cash origin for the			
-	cambe of the galax that has not a cample over these true. Believe and only cambo and states to the conservation of an angle cambo later to the conservation of a part and provide to			
148	namber efficienzat al an inderstaat onder has the excita templometer dater fact, foal is used where he fight make excernal are grow tables norther	-		
	satise of to stary lighter hadings dier for to adde satisfies and light for.	-		
-	per againe fragme, north, and day observe are he reacted damp in semicirceller.	-		
	no.e. signer have rent, and the minimum term			
-	ing regular to provide work, and dependences for canonal angular associations.			
	and the	-		

View/Open

data_grobal-noble-gases.tsr (1 120Mb)
ISO19115-2.xml (338.3Kb)
Field_names.pdi (29.75Kb)
Global_Hammeetal2019.mat (400.0Kb)
Readme_Hammeetal2019.txt

(20.55Kb)

Dataset_description.pdf (197.0Kb)

Date

2018-08-24

Author

Hamme, Roberta C. ⊘ Jenkins, William J. ⊘

Citable URI

https://hdl.handle.net/1912/10538

Date Created

2018-08-13

Location

North Atlantic

Eastern Tropical Pacific - Transect from Peru to Tahiti Oxygen minium zone; East Pacific Rise Global oceans westlimit: -159.9952; southlimit: -68.1081; eastlimit: 178.9985; northlimit: 78.9988

DOI

10.1575/1912/bco-dmo.744563

Inert gases dissolved in the ocean are powerful tracers of the impact of physical processes on gases, particularly air-sea gas exchange (by both diffusive and bubblemeditated processes), temperature change, atmospheric pressure variation, mixing between different water masses, and ice processes. We have compiled a global ocean database of dissolved neon, argon, and krypton measurements, supplemented by helium, xenon, and nitrogen/argon (N2/Ar) ratios in some locations. Samples were collected on board multiple research cruises spanning the period 1999 through 2016 and analyzed by mass spectrometry at four different shore-based laboratories (University of Victoria, Woods Hole Oceanographic Institution, University of Washington, and Scripps Institution of Oceanography). For a complete list of measurements, refer to the supplemental document 'Field_names.pdf', and a full dataset description is included in

Static ISO record in the Archive w. old information







old definition of "photosynthetic rate"

The gmx:Anchor tag

<gco:CharacterString>photosynthetic rate</gco:CharacterString>

becomes

<gmx:Anchor xlink:href="{URL goes here}">photosynthetic rate</gmx:Anchor>

Things instead of Strings

What is meant by a string

at a single point in time

may change in the future.

Photosynthetic Rate as a "Thing" It has a label. It has a definition. It can have synonyms. It can be related URI to other things.

Date ()

2016-06-20 12:31:50.0

http://vocab.nerc.ac.uk/collection/P21/current/MS8092/ Identifier () SDN:P21::MS8092 Preferred label (en) photosynthesis Alternative label () Photosynthesis is a process by which plants and other organisms use light energy and chlo Definition (en) compounds. Photosynthesis is one of the mechanisms responsible for Primary Production. measurement. Version Info () 3 Has Current Version http://vocab.nerc.ac.uk/collection/P21/current/MS8092/3/ http://vocab.nerc.ac.uk/collection/P21/current/MS8092/2/ Has Version Has Version http://vocab.nerc.ac.uk/collection/P21/current/MS8092/1/ PAV Version () 3 PAV Authored On () 2016-06-20 12:31:50.0 Deprecated() false http://vocab.nerc.ac.uk/collection/P21/current/MS8687/ Broader http://vocab.nerc.ac.uk/collection/P21/current/MS3484/ Narrower http://vocab.nerc.ac.uk/collection/P21/current/MS10803/ Narrower http://vocab.nerc.ac.uk/collection/P21/current/MS4551/ Narrower http://vocab.nerc.ac.uk/collection/P21/current/MS780/ Narrower http://vocab.nerc.ac.uk/collection/P21/current/MS11730/ Related

"photosynthetic rate"

"Photosynthesis is a process..."

"photosynthesis"

Photosynthetic Rate as a "Thing"

All important pieces of information

related to a dataset get an identifier.

http://lod.bco-dmo.org/id/parameter/648663

W3C Best Practices for Data on the Web

https://www.w3.org/TR/dwbp/#bp-summary

Best Practice 1: Provide metadata Best Practice 2: Provide descriptive metadata Best Practice 3: Provide structural metadata Best Practice 4: Provide data license information Best Practice 5: Provide data provenance information Best Practice 6: Provide data quality information Best Practice 7: Provide a version indicator Best Practice 8: Provide version history Best Practice 9: Use persistent URIs as identifiers of datasets Best Practice 10: Use persistent URIs as identifiers within datasets Best Practice 11: Assign URIs to dataset versions and series Best Practice 12: Use machine-readable standardized data formats Best Practice 13: Use locale-neutral data representations Best Practice 14: Provide data in multiple formats Best Practice 15: Reuse vocabularies, preferably standardized ones Best Practice 16: Choose the right formalization level Best Practice 17: Provide bulk download Best Practice 18: Provide Subsets for Large Datasets

Best Practice 19: Use content negotiation for serving data available in multiple formats Best Practice 20: Provide real-time access Best Practice 21: Provide data up to date Best Practice 22: Provide an explanation for data that is not available Best Practice 23: Make data available through an API Best Practice 24: Use Web Standards as the foundation of APIs Best Practice 25: Provide complete documentation for your API Best Practice 26: Avoid Breaking Changes to Your API Best Practice 27: Preserve identifiers Best Practice 28: Assess dataset coverage Best Practice 29: Gather feedback from data consumers Best Practice 30: Make feedback available Best Practice 31: Enrich data by generating new data Best Practice 32: Provide Complementary Presentations Best Practice 33: Provide Feedback to the Original Publisher Best Practice 34: Follow Licensing Terms Best Practice 35: Cite the Original Publication

Best Practice #10:

Use persistent URIs as identifiers *within* datasets

https://www.w3.org/TR/dwbp/#identifiersWithinDatasets

Use persistent URIs as identifiers within datasets

https://www.w3.org/TR/dwbp/#identifiersWithinDatasets



http://lod.bco-dmo.org/id/parameter/648663

In a browser https://www.bco-dmo.org/parameter/648663

Supports Content Negotiation:

GET http://lod.bco-dmo.org/id/parameter/648663 Accept: text/html \rightarrow HTML Accept: application/xml \rightarrow RDF/XML

Supports URL File Extensions:

GET http://lod.bco-dmo.org/id/parameter/648663.xml GET http://lod.bco-dmo.org/id/parameter/648663.rdf





 $\rightarrow RDF/XML$ $\rightarrow RDF/XML$

The gmx: Anchor tag with URIs

<gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/648663.rdf">...

NOAA ISO 19115-2 snapshot of Photosynthetic Rate

▼<qmd:keyword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1617.rdf" xlink:title="Parameter" xlink:actuate="onReguest">het oxygen production</gmx:Anchor> </gmd:keyword> ▼<qmd:keyword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1620.rdf" xlink:title="Parameter" xlink:actuate="onReguest">oxygen respiration rates from in situ incubation</gmx:Anchor> </gmd:keyword> ▼<qmd:kevword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/648663.rdf" xlink:title="Parameter" xlink:actuate="onRequest">photosynthetic rate</gmx:Anchor> </gmd:kevword> ▼<qmd:kevword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1073.rdf" xlink:title="Parameter" xlink:actuate="onReguest">no standard parameter/gmx:Anchor> </gmd:keyword> ▼<gmd:type> <qmd:MD KeywordTypeCode codeList="http://www.isotc211.org/2005/resources/Codelist/amxCodelists.xml#MD KeywordTypeCode"</pre> codeListValue="featureType">featureType</qmd:MD KeywordTypeCode> </gmd:type> v<qmd:thesaurusName> ▼<gmd:CI Citation> v<gmd:title> <gco:CharacterString>BCO-DMO Standard Parameters</gco:CharacterString>

NOAA ISO 19115-2 snapshot of Photosynthetic Rate

▼<qmd:keyword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1617.rdf" xlink:title="Parameter" xlink:actuate="onReguest">het oxygen production</gmx:Anchor> </gmd:keyword> ▼<qmd:keyword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1620.rdf" xlink:title="Parameter" xlink:actuate="onReguest">oxygen respiration rates from in situ incubation</gmx:Anchor> </gmd:keyword> ▼<qmd:kevword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/648663.rdf" xlink:title="Parameter" xlink:actuate="onRefuest">photosynthetic rate</gmx:Anchor> </amd:keyword> ▼<qmd:kevword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1073.rdf" xlink:title="Parameter" xlink:actuate="onReguest">no standard parameter/gmx:Anchor> </gmd:keyword> ▼<gmd:type> <qmd:MD KeywordTypeCode codeList="http://www.isotc211.org/2005/resources/Codelist/amxCodelists.xml#MD KeywordTypeCode"</pre> codeListValue="featureType">featureType</qmd:MD KeywordTypeCode> </gmd:type> v<gmd:thesaurusName> ▼<gmd:CI Citation> v<gmd:title> <gco:CharacterString>BCO-DMO Standard Parameters</gco:CharacterString>

NOAA ISO 19115-2 snapshot of Photosynthetic Rate

▼<qmd:keyword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1617.rdf" xlink:title="Parameter" xlink:actuate="onReguest">het oxygen production</gmx:Anchor> </gmd:keyword> ▼<qmd:keyword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1620.rdf" xlink:title="Parameter" xlink:actuate="onReguest">oxygen respiration rates from in situ incubation</gmx:Anchor> </gmd:keyword> ▼<qmd:kevword> <gmx:Anchor xlink:hrg="http://lod.bco-dmo.org/id/parameter/648663.rdf" xlnk:title="Parameter" xlink:actuate="onReguest">photosynthetic rate</gmx:Anchor> </amd:keyword> ▼<gmd:kevword> <gmx:Anchor xlink:href="http://lod.bco-dmo.org/id/parameter/1073.rdf" xlink:title="Parameter" xlink:actuate="onReguest">no standard parameter</gmx:Anchor> </gmd:keyword> ▼<gmd:type> <qmd:MD KeywordTypeCode codeList="http://www.isotc211.org/2005/resources/Codelist/amxCodelists.xml#MD KeywordTypeCode"</pre> codeListValue="featureType">featureType</qmd:MD KeywordTypeCode> </gmd:type> v<qmd:thesaurusName> ▼<gmd:CI Citation> v<gmd:title> <gco:CharacterString>BCO-DMO Standard Parameters</gco:CharacterString>

BCO-DMO Strategy for ISO 19115-2 XML

- Use <gmx:Anchor> instead of <gco:CharacterString> (Thing) (String)
- For humans, we use a file extension URL for readability (.rdf)
- For machines, gmx:Anchor requires xlink:href URL resolve to XML (Content-Type → application/rdf+xml)

http://lod.bco-dmo.org/id/parameter/648663.rdf

Questions?

Example ISO record at BCO-DMO: https://www.bco-dmo.org/dataset/648416/iso/NOAA_ISO-19115-2.xml