



# ***Adoption of Data Citation Outcomes by BCO-DMO***

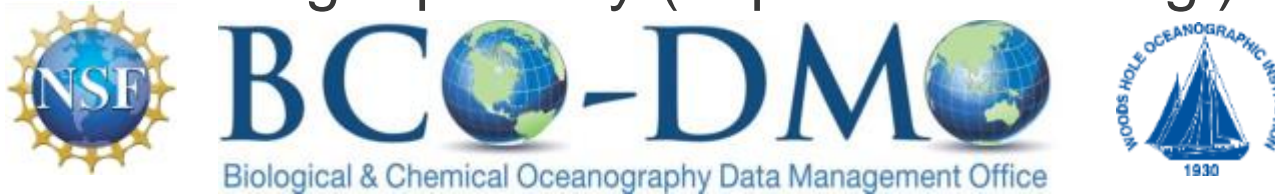
**Cynthia Chandler, Adam Shepherd, David Bassendine  
Biological and Chemical Oceanography Data Management Office  
Woods Hole Oceanographic Institution and Blue Dot Labs**

**research data sharing without barriers**  
**[rd-alliance.org](http://rd-alliance.org)**

# *A story of success enabled by RDA*

2

- An existing repository (<http://bco-dmo.org/>)





- Research data curation for a decade
- Faced with new challenges, but no new funding
- e.g. data publication practices to support citation
- Used the outcomes from the RDA Data Citation Working Group to improve data publication and citation services

<https://www.rd-alliance.org/groups/data-citation-wg.html>

- BCO-DMO is a thematic, domain-specific repository funded by NSF Ocean Sciences and Polar Programs
- BCO-DMO curated data are
  - Served: <http://bco-dmo.org> (URLs, URIs)
  - Published: at an Institutional Repository (CrossRef DOI) <http://dx.doi.org/10.1575/1912/4847>
  - Archived: at NCEI, a US National Data Center <http://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0078575>

## Archival Copy

for Linked Data URI: <http://lod.bco-dmo.org/id/dataset/3046>

Version Date	Archive	Persistent Identifier	Date Assigned
2007-05-24	National Oceanographic Data Center (NODC)	0078575 	2011-10-26
2007-05-24	Marine Biological Laboratory/Woods Hole Oceanographic Institution Library (MBLWHOI DLA)	<a href="http://dx.doi.org/10.1575/1912/4847">10.1575/1912/4847</a> 	2011-10-12

- **DC goals:** to create identification mechanisms that:
  - allow us to identify and cite arbitrary views of data, from a single record to an entire data set in a precise, machine-actionable manner
  - allow us to cite and retrieve that data as it existed at a certain point in time, whether the database is static or highly dynamic
- **DC outcomes:** 14 recommendations and associated documentation
  - ensuring that data are stored in a versioned and timestamped manner
  - identifying data sets by storing and assigning persistent identifiers (PIDs) to timestamped queries that can be re-executed against the timestamped data store

<https://www.rd-alliance.org/groups/data-citation-wg.html>

- Preparing the Data and the Query Store
  - R1 - Data Versioning
  - R2 - Timestamping
  - R3 - Query Store Facilities
- Persistently Identifying Specific Data Sets
  - R4 - Query Uniqueness
  - R5 - Stable Sorting
  - R6 - Result Set Verification
  - R7 - Query Timestamping
  - R8 - Query PID
  - R9 - Store the Query
  - R10 - Automated Citation Texts
- Resolving PIDs and Retrieving the Data
  - R11 - Landing Page
  - R12 - Machine Actionability
- Upon modifications to the Data Infrastructure
  - R13 - Technology Migration
  - R14 - Migration Verification

»» Data Versioning: For retrieving earlier states of datasets the data need to be versioned. Markers shall indicate inserts, updates and deletes of data in the database.

»» Data Timestamping: Ensure that operations on data are timestamped, i.e. any additions, deletions are marked with a timestamp.


»» Data Identification: The data used shall be identified via a PID pointing to a time-stamped query, resolving to a landing page.

Oct 2015 version w/ 14 recommendations

DC WG chairs: Andreas Rauber, Ari Asmi, Dieter van Uytvanck

# BCO-DMO Dataset Landing Page

6



Biological & Chemical Oceanography Data Management Office

[HOME](#)[DATA](#)[RESOURCES](#)[ABOUT US](#)

**DATABASE**


Programs	36
Projects	649
Deployments	2300
Datasets	7892
Instruments	395
Parameters	1366
People	2026
Affiliations	472
Funding	79
Awards	1304

**Dataset: larval krill pigments**

[Get Data](#)[Map It](#)

<b>Project:</b> <a href="#">U.S. GLOBEC Southern Ocean (SOGLOBEC)</a>
<b>Principal Investigator:</b> <a href="#">Dr Robin Ross</a> (University of California-Santa Barbara, UCSB) <a href="#">Dr Langdon Quetin</a> (University of California-Santa Barbara, UCSB)
<b>Contact:</b> <a href="#">Dr Robin Ross</a> (University of California-Santa Barbara, UCSB)
<b>BCO-DMO Data Manager:</b> <a href="#">Nancy Copley</a> (Woods Hole Oceanographic Institution, WHOI BCO-DMO)
<b>Validated:</b> Yes
<b>Data version:</b> 2010-02-03
<b>Version Date:</b> 02/03/2010
<b>Data URL:</b> <a href="http://www.bco-dmo.org/dataset/3300/data">http://www.bco-dmo.org/dataset/3300/data</a> <a href="#">↗</a>
<b>Current State:</b> Final no updates expected

**GEOSPATIAL ACCESS**



[Expand/Collapse All](#)

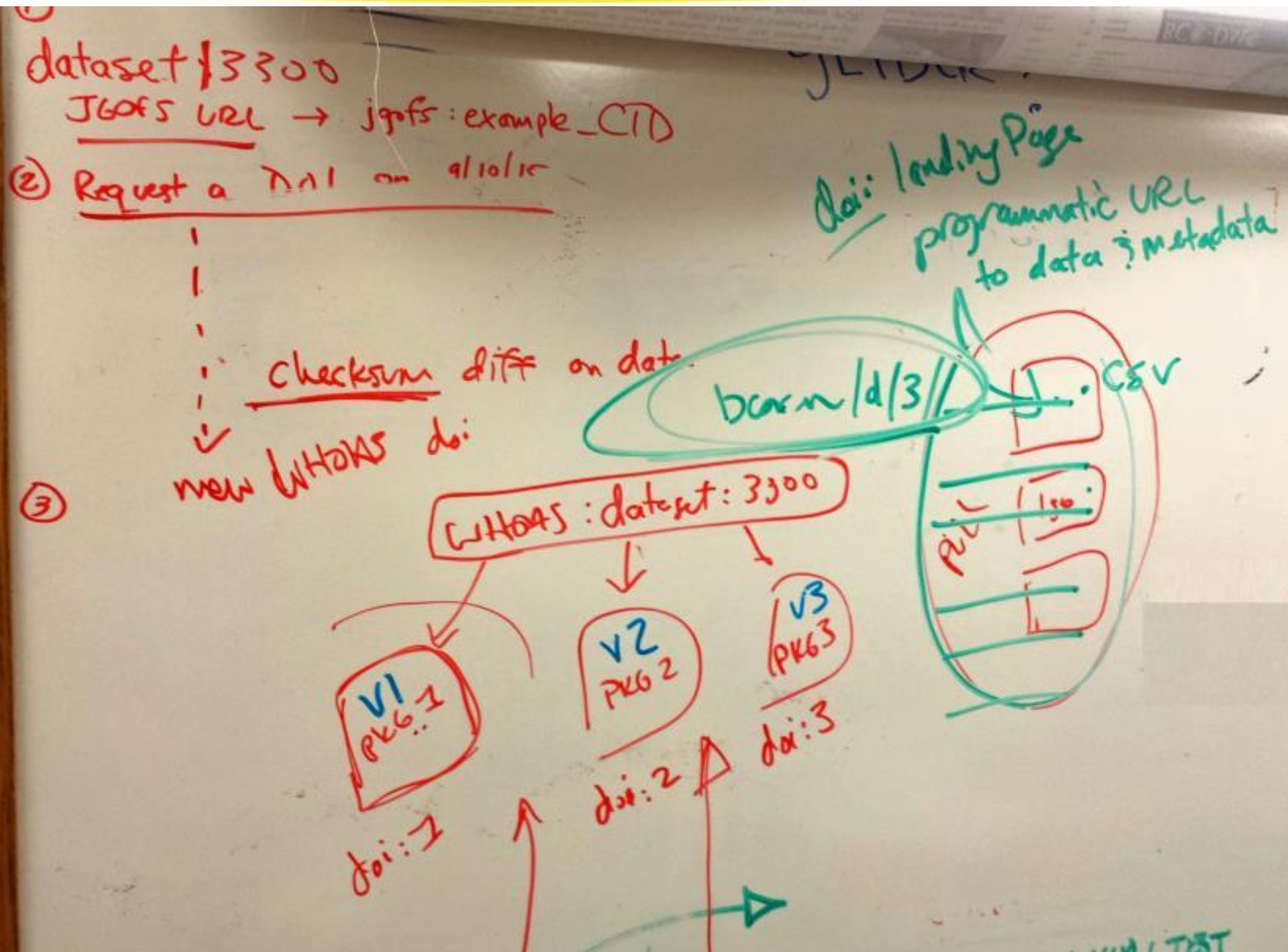
▼ [Archival Copy](#)

Archive Location	Accession ID	Archive Copy URL
National Oceanographic Data Center (NODC)	0112635	<a href="http://accession.nodc.noaa.gov/0112635">http://accession.nodc.noaa.gov/0112635</a>



# Initial Architecture Design Considerations (Jan 2016)

7



# Modified Architecture (March 2016)

8

## BCO-DMO DATA PUBLICATION & CITATION

28 MAR 2016

Archive Record: LWHORS [mid = 567890]

DFD 90001 [ISO] download: DOI  
doi: 10.175/1915/bco-dmo-567890

DFD 90002 [DATA 1]

DFD 90003 [DATA N]

LWHORS

[ISO]

[1]

[N]

QUESTION:

supporting docs  
are recorded as DFD  
at LWHORS for  
reuse w/ NCET

DATA FILE DESCRIPTOR [DFD]

OSPREY CONTENT TYPE

checksum: text  
checksum algorithm: controlled vocabulary  
filesize\_in\_bytes: integer  
resolvable\_url: URL  
dataset: Dataset or Dataset Deployment  
type: controlled vocabulary [Data File Type]

DATA File Type

name: text (ex: Tab Separated Values, ISO 14175-2)  
data\_one\_identifier: text  
data\_one\_type: controlled vocabulary  
mime\_type: text (text/csv, application/xml)

Archive Record: NCET [mid = 67890]

DFD 90001  
DFD 90002  
DFD 90003

Serialized from LWHORS,  
Stored in WAF (HTTPS)  
SHA 384 Manifest & Summary

Archive Record: DATAONE [mid = 789012]

DFD 90001 [ISO] download: DOI  
Accession # at NCET  
DOI: 10.175/1915/bco-dmo-567890

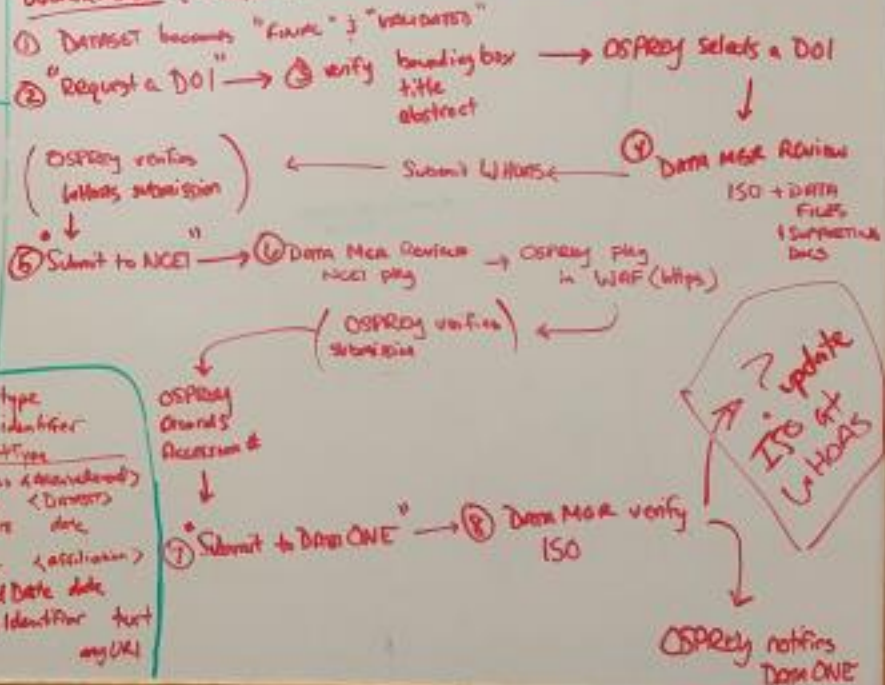
DFD 90002

DFD 90003

(stored locally  
in OSPREY)  
or update  
LWHORS?

(Serialized from  
LWHORS)

WORKFLOW (8 steps for new DATASET VICEPRES)



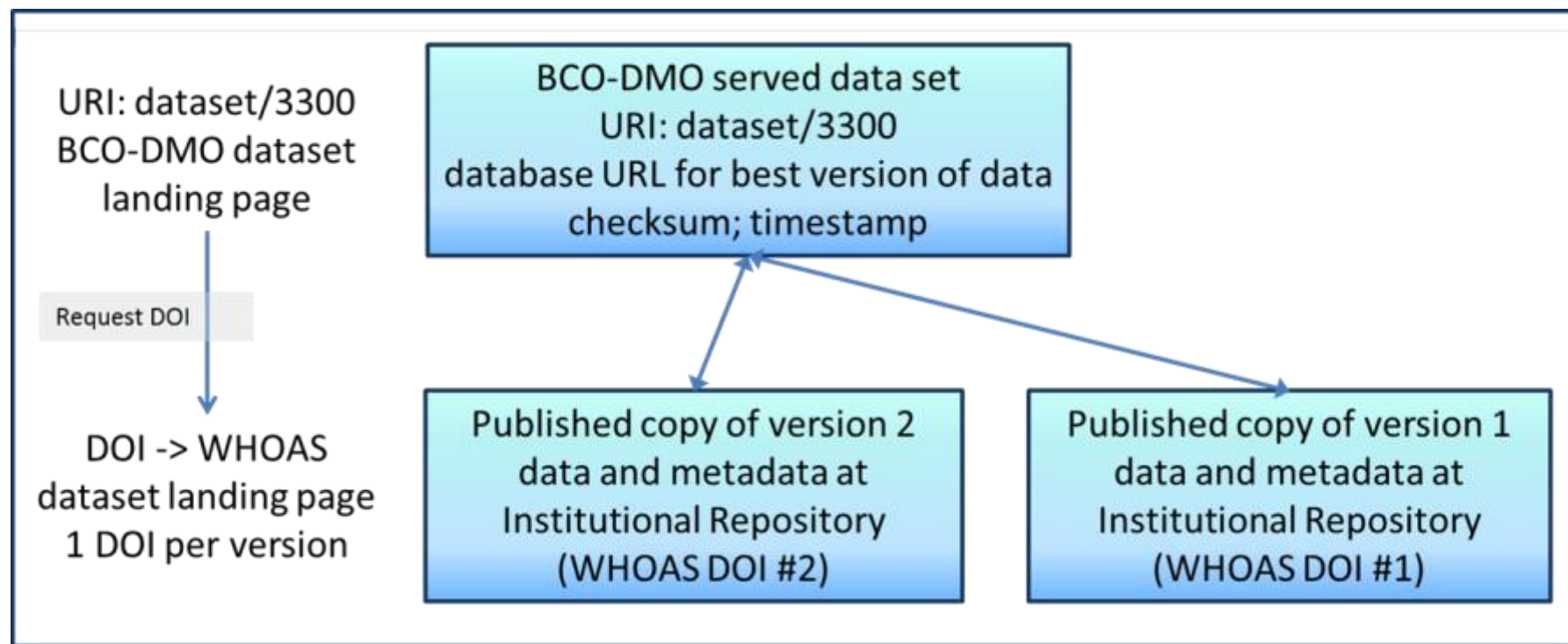
<DATASET>	archived By	<Archive Record>
Archive Record	has Bitstream	
Data File Descriptor		
checksum	string	
checksum algorithm	string	
filesize_in_bytes	int	
resolvable_url	any URL	
dataset	any URL	
type		
mime_type		

<Data File Type>	dfi type	dfi identifier	subset type
archive	abstract & document		
dataset	abstract		
version	version	date	
title	title		
creator	creator		
archived date	date		
resolvable identifier	text		
url	any URL		



# BCO-DMO Data Publication System Components

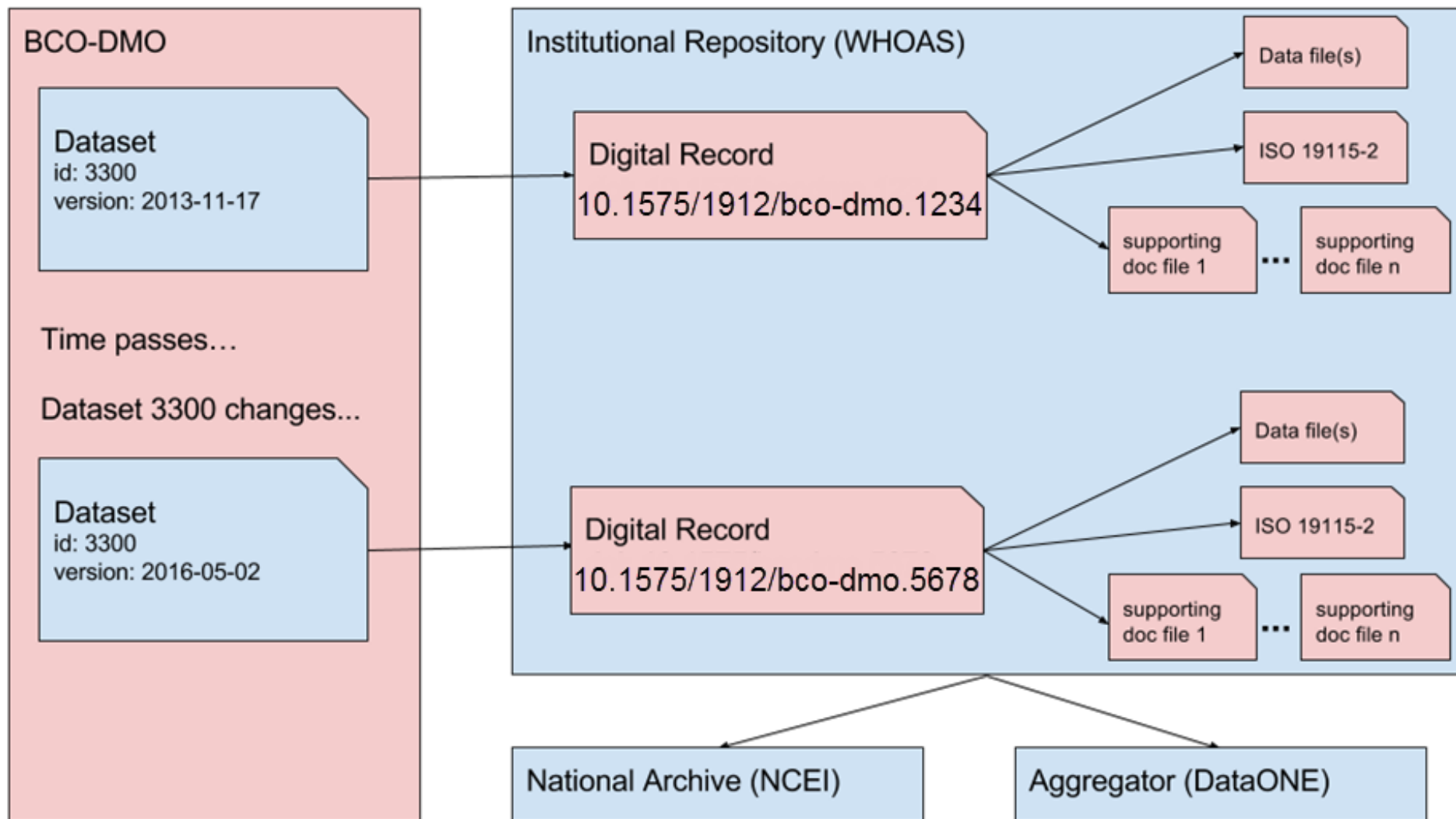
9



BCO-DMO publishes data to WHOAS and a DOI is assigned.  
The BCO-DMO architecture now supports data versioning.


# BCO-DMO Data Citation System Components

10



# BCO-DMO Data Set Landing Page

11



Biological & Chemical Oceanography Data Management Office

DATA RESOURCES ABOUT US

Enter search terms

DATABASE

- New Entry
- Programs 36
- Projects 696
- Deployments 2360
- Platforms 486
- Datasets 8079
- Instruments 408

### Dataset: Cellular element quotas: Si in Synechococcus cells

[Get Data](#) [Map It](#)

**Project:** Understanding the Role of Picocyanobacteria in the Marine Silicate Cycle (Si\_in\_Syn)

**Principal Investigator:** Dr Benjamin Twining (Bigelow Laboratory for Ocean Sciences, Bigelow)

**Contact:** Daniel Ohnemius (Bigelow Laboratory for Ocean Sciences, Bigelow)

**Data version:** Final

**Version Date:** 05/06/2016

▼ [Archival Copy](#)

## ▼ Archival Copy

Version Date	Archive	Persistent Identifier	Date Assigned
2016-05-06	<a href="#">Marine Biological Laboratory/Woods Hole Oceanographic Institution Library (MBLWHOI DLA)</a>	<a href="https://doi.org/10.1575/1912/bco-dmo.651474">10.1575/1912/bco-dmo.651474</a> 	2016-07-08

## GEOSPATIAL ACCESS



## References:

Twining, B.S., et al. Metal contents of phytoplankton and labile particulate material in the North Atlantic Ocean. *Prog. Oceanogr.* (2015)  
<http://dx.doi.org/10.1016/j.pocean.2015.07.001>  
[https://www.researchgate.net/publication/282626294\\_Metal\\_contents\\_of\\_phytoplankton\\_and\\_labile\\_particulate\\_material\\_in\\_the\\_North\\_Atlantic\\_Ocean](https://www.researchgate.net/publication/282626294_Metal_contents_of_phytoplankton_and_labile_particulate_material_in_the_North_Atlantic_Ocean)

▼ [More Information about this dataset](#)

Search



- ☒ Search WHOAS  
☐ This Collection

BROWSE

All of WHOAS

Communities & Collections

# Element quotas of individual *Synechococcus* cells collected during Bermuda Atlantic Time-series Study (BATS) cruises aboard the R/V Atlantic Explorer between dates 2012-07-11 and 2013-10-13

**dc.description.sponsorship** NSF Ocean Sciences (NSF OCE) OCE-1131139, NSF Ocean Sciences (NSF OCE) OCE-1335012, NSF Ocean Sciences (NSF OCE) OCE-1131046

**dc.publisher** Biological and Chemical Oceanography Data Management Office (BCO-DMO). Contact: bco-dmo-data@whoi.edu




**dc.relation** <http://lod.bco-dmo.org/id/dataset/644840>

By Issue Date

Authors

Titles

View Item

-  cellular-element-quotas-si-synechococcus-cells.tsv (10.09Kb)  
 115-2.xml (88.65Kb)  
 Names.pdf (12.10Kb)

31.6691

**DOI**  
[10.1575/1912/bco-dmo.651474](https://doi.org/10.1575/1912/bco-dmo.651474)


**Keyword**

SXRF; Synchrotron radiation X-Ray Fluorescence; *Synechococcus*; si



# BCO-DMO Data Set Landing Page

13



Biological & Chemical Oceanography Data Management Office

[DATA](#) [RESOURCES](#) [ABOUT US](#)

**DATABASE**

- New Entry
- Programs 36
- Projects 696
- Deployments 2360
- Platforms 486
- Datasets 8079
- Instruments 408
- Parameters 1373

### Dataset: Cellular element quotas: Si in Synechococcus cells

[Get Data](#) [Map It](#)

**Project:** Understanding the Role of Picocyanobacteria in the Marine Silicate Cycle (Si\_in\_Syn)

**Principal Investigator:** Dr Benjamin Twining (Bigelow Laboratory for Ocean Sciences, Bigelow)

**Contact:** Daniel Ohnemus (Bigelow Laboratory for Ocean Sciences, Bigelow)

**Data version:** Final

**Version Date:** 05/06/2016

▼ **Archival Copy**

Version	Archive	Persistent Identifier	Date
---------	---------	-----------------------	------

## References:

Twining, B.S., et al. Metal contents of phytoplankton and labile particulate material in the North Atlantic Ocean. Prog. Oceanogr. (2015)  
<http://dx.doi.org/10.1016/j.pocean.2015.07.001>  
[https://www.researchgate.net/publication/282626294\\_Metal\\_contents\\_of\\_phytoplankton\\_and\\_labile\\_particulate\\_material\\_in\\_the\\_North\\_Atlantic\\_Ocean](https://www.researchgate.net/publication/282626294_Metal_contents_of_phytoplankton_and_labile_particulate_material_in_the_North_Atlantic_Ocean)

are also provided.

## GEOSPATIAL ACCESS






## References:


Twining, B.S., et al. Metal contents of phytoplankton and labile particulate material in the North Atlantic Ocean. Prog. Oceanogr. (2015)  
<http://dx.doi.org/10.1016/j.pocean.2015.07.001>  
[https://www.researchgate.net/publication/282626294\\_Metal\\_contents\\_of\\_phytoplankton\\_and\\_labile\\_particulate\\_material\\_in\\_the\\_North\\_Atlantic\\_Ocean](https://www.researchgate.net/publication/282626294_Metal_contents_of_phytoplankton_and_labile_particulate_material_in_the_North_Atlantic_Ocean)

▼ [More information about this dataset](#)

**ScienceDirect**

JournalsBooksSign in ?

 Download PDF  Add to online library  Export   Advanced search



**Progress in Oceanography**  
Volume 137, Part A, September 2015, Pages 261–283



## Metal contents of phytoplankton and labile particulate material in the North Atlantic Ocean

Benjamin S. Twining<sup>a</sup>, Sara Rauschenberg<sup>a</sup>, Peter L. Morton<sup>b, 1</sup>, Stefan Vogt<sup>c</sup>

[Show more](#)

[doi:10.1016/j.pocean.2015.07.001](https://doi.org/10.1016/j.pocean.2015.07.001)[Get rights and content](#)

### Highlights

- First basin-wide measurements of plankton metal quotas in the N. Atlantic Ocean.
- Fe and Mn quotas significantly higher on western side of section.
- Cu and Ni quotas significantly elevated on eastern side of section.
- Evidence for Al scavenging by biogenic silica.
- Dissolved ratios not an accurate measure of cellular Fe quotas.

### Recommended articles

**Evaluation of approaches to estimate biogenic ...**  
2015, Marine Chemistry [more](#)

**High-frequency sea level oscillations in the Me...**  
2015, Progress in Oceanography [more](#)

**Composition of metals in suspended particulate...**  
2016, Regional Studies in Marine Science [more](#)

[View more articles »](#)

### Citing articles (2)

### Related book content

# New Capabilities ... BCO-DMO becoming a DataONE Member Node

15

<https://search.dataone.org/>

The screenshot displays the DataONE search interface. At the top, the DataONE logo is on the left, and navigation links (About, News, Participate, Resources, Education, Data) are in the center. A search bar on the right includes a 'DATAONE SEARCH:' label, a 'Search' button, a 'Summary' button, a 'Jump to:' dropdown set to 'DOI or ID', a 'Go' button, and 'Sign In' or 'Sign up' links.

Below the navigation bar, a teal bar contains a 'Clear all filters' link. The main content area is divided into three sections:

- Search ?**: A search bar with the placeholder 'Search phrase' and a magnifying glass icon.
- My Search**: A box containing the text 'BCO-DMO' with a close button (X).
- Filter by:**: A list of filter categories: Data attribute, Data files, Member Node, Creator, Year, Identifier, and Taxon.

The central section displays 'Datasets 1 to 25 of 149'. It includes a pagination bar with numbers 1, 2, 3, ..., 6, and a 'Next' button. A 'Sort by' dropdown is set to 'Most recent'. Three dataset entries are visible:

- Dr David Siegel and Dr Kenneth O. Buesseler. 2009. urn:nod:mTestBCODMO. <http://lod.bco-dmo.org/id/dataset-file/555907>.
- Dr Carl Lamborg and Dr Kenneth O. Buesseler. 2009. urn:nod:mTestBCODMO. <http://lod.bco-dmo.org/id/dataset-file/555902>.
- Dr Uta Passow. 2014. urn:nod:mTestBCODMO. <http://lod.bco-dmo.org/id/dataset-file/555897>.

Each entry is accompanied by icons for information, a document, and a location pin. At the bottom of the dataset list, there is a logo for 'DOC/NOAA/NESDIS/NCEI > National Centers for Environmental Information, NESDIS, NOAA, U.S.'.

On the right side of the interface is a map of the United States and surrounding regions. The map shows state boundaries and is overlaid with a grid of numbered blue squares. The numbers visible are 2, 4, 19, 1, 25, and 1. The map includes a 'Hide Map' button at the top, zoom controls (+, -), and a legend for 'Satellite' and 'Terrain'. At the bottom of the map, it says 'Google Map data ©2016 Google, INEGI 1000 km Terms of Use'.

# BCO-DMO Data Set Citation (by 19 Sep 2016)

16

The screenshot shows the BCO-DMO website interface. At the top is the BCO-DMO logo and navigation buttons for DATA, RESOURCES, and ABOUT US. A search bar is on the right. On the left, a 'DATABASE' sidebar lists 'New Entry', 'Programs' (36), 'Projects' (696), and 'Deployments' (2360). The main content area displays the dataset title 'Dataset: Cellular element quotas: Si in Synechococcus cells' with 'Get Data' and 'Map It' buttons. Below this, it lists the 'Project: Understanding the Role of Picocyanobacteria in the Marine Silicate Cycle (Si\_in\_Syn)' and the 'Principal Investigator: Dr Benjamin Twining (Bigelow Laboratory for Ocean Sciences, Bigelow)'. A 'Citation' button with a Creative Commons icon is highlighted with an arrow.

## Suggested citation:

Twining, B. (2016) "Element Quotas of Individual Synechococcus Cells Collected During Bermuda Atlantic Time-Series Study (BATS) Cruises Aboard the R/V Atlantic Explorer Between Dates 2012-07-11 and 2013-10-13". Dataset published by Biological and Chemical Oceanography Data Management Office (BCO-DMO) doi:10.1575/1912/bco-dmo.651474

This screenshot shows the lower portion of the BCO-DMO dataset page. On the left, a sidebar lists 'Funding' (83), 'Awards' (1378), 'Contact Status' (270), and 'Archive' (187). Below this is a 'GEOSPATIAL ACCESS' section with a map of the North Atlantic Ocean showing sampling locations. The main content area includes a 'Library (MBLWHOI DLA)' dropdown, a 'Description' section with a 'Brief Description: element quotas of individual Synechococcus cells', and a 'References' section listing a publication by Twining, B.S., et al. (2015) with its DOI and ResearchGate link. A 'More Information about this dataset' link is at the bottom.



# Thank you ...

17

- To the Data Citation Working Group for their efforts  
<https://www.rd-alliance.org/groups/data-citation-wg.html>
- RDA US for funding this adoption project
- TIMELINE:
  - Redesign/protoype completed by 1 June
  - New citation recommendation by 19 Sep 2016
  - Report out at RDA P8 (Denver, CO) September 2016
  - Production version by 1 December 2016

Cyndy Chandler      @cynDC42  
ORCID: 0000-0003-2129-1647

@bcodmo  
cchandler@whoi.edu

- Removed these to reduce talk to 15 minutes

## ■ Evaluation

- Evaluate recommendations (done December 2015)
- Try implementation in existing BCO-DMO architecture (work began 4 April 2016)

## ■ Trial

- BCO-DMO: R1-11 fit well with current architecture; R12 doable; test as part of DataONE node membership; R13-14 are consistent with Linked Data approach to data publication and sharing

**NOTE: adoption grant received from RDA US (April 2016)**

# New capability (implemented)

procedure: when a BCO-DMO data set is updated ...

- A copy of the previous version is preserved
- Request a DOI for the new version of data
- Publish data, and create new landing page for new version of data, with new DOI assigned
- BCO-DMO database has links to all versions of the data (archived and published)
- Both archive and published dataset landing pages have links back to best version of full dataset at BCO-DMO
- BCO-DMO data set landing page displays links to all archived and published versions



- Extended description of recommendations  
**Identification of Reproducible Subsets for Data Citation,  
Sharing and Re-Use**

Andreas Rauber  
Vienna University of Technology, Austria  
rauber@ifs.tuwien.ac.at

Dieter van Uytvanck  
CLARIN ERIC, Utrecht, Netherlands  
dieter@clarin.eu

Ari Asmi  
University of Helsinki, Finland  
ari.asmi@helsinki.fi

Stefan Pröll  
SBA Research, Vienna, Austria  
sproell@sba-research.org

- Altman and Crosas. 2013. “Evolution of Data Citation ...”
- CODATA-ICSTI 2013. “Out of cite, out of mind”
- FORCE11 <https://www.force11.org/about/mission-and-guiding-principles>
- R. E. Duerr, et al. “On the utility of identification schemes for digital earth science data”, ESI, 2011.