RESEARCH DATA ALLIANCE

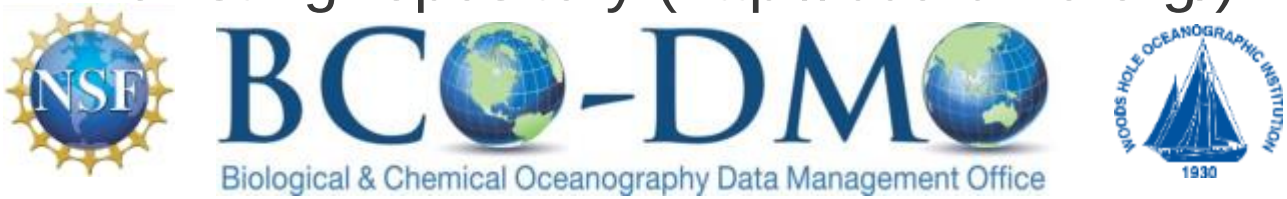
Adoption of Data Citation Outcomes by BCO-DMO

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

A story of success enabled by RDA

- 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 Curated Data

- 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)	10.1575/1912/4847	2011-10-12

RDA Data Citation (DC) *of evolving data*

- **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

RDA Data Citation WG Recommendations

- 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

BCO-DMO
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

GEOSPATIAL ACCESS

Dataset: larval krill pigments

Get Data | Map It

Project: U.S. GLOBEC Southern Ocean (SOGLOBEC)

Principal Investigator: Dr Robin Ross (University of California-Santa Barbara, UCSB)
Dr Langdon Quetin (University of California-Santa Barbara, UCSB)

Contact: Dr Robin Ross (University of California-Santa Barbara, UCSB)

BCO-DMO Data Manager: Nancy Copley (Woods Hole Oceanographic Institution, WHOI BCO-DMO)

Validated: Yes

Data version: 2010-02-03

Version Date: 02/03/2010

Data URL: http://www.bco-dmo.org/dataset/3300/data

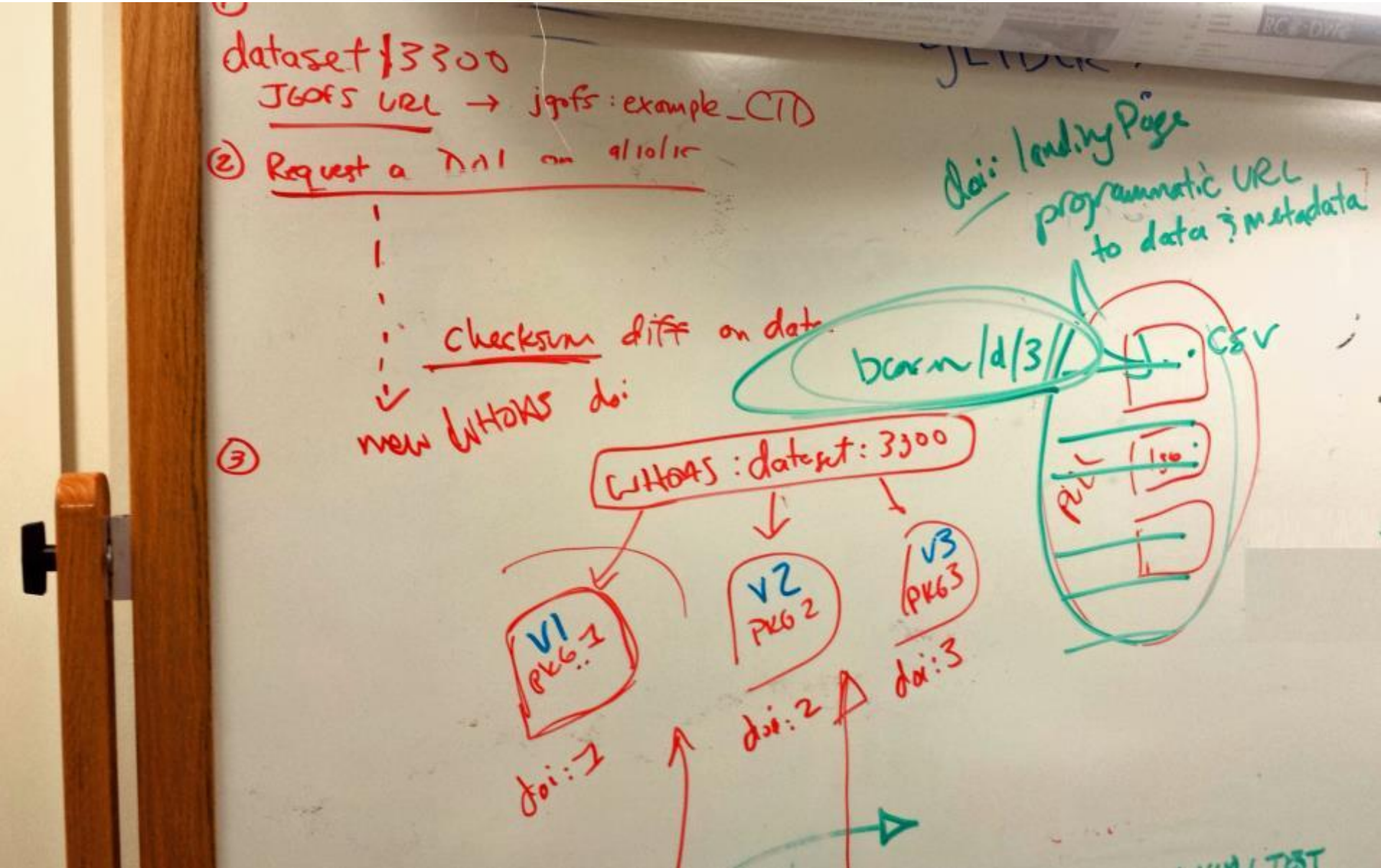
Current State: Final no updates expected

Expand/Collapse All

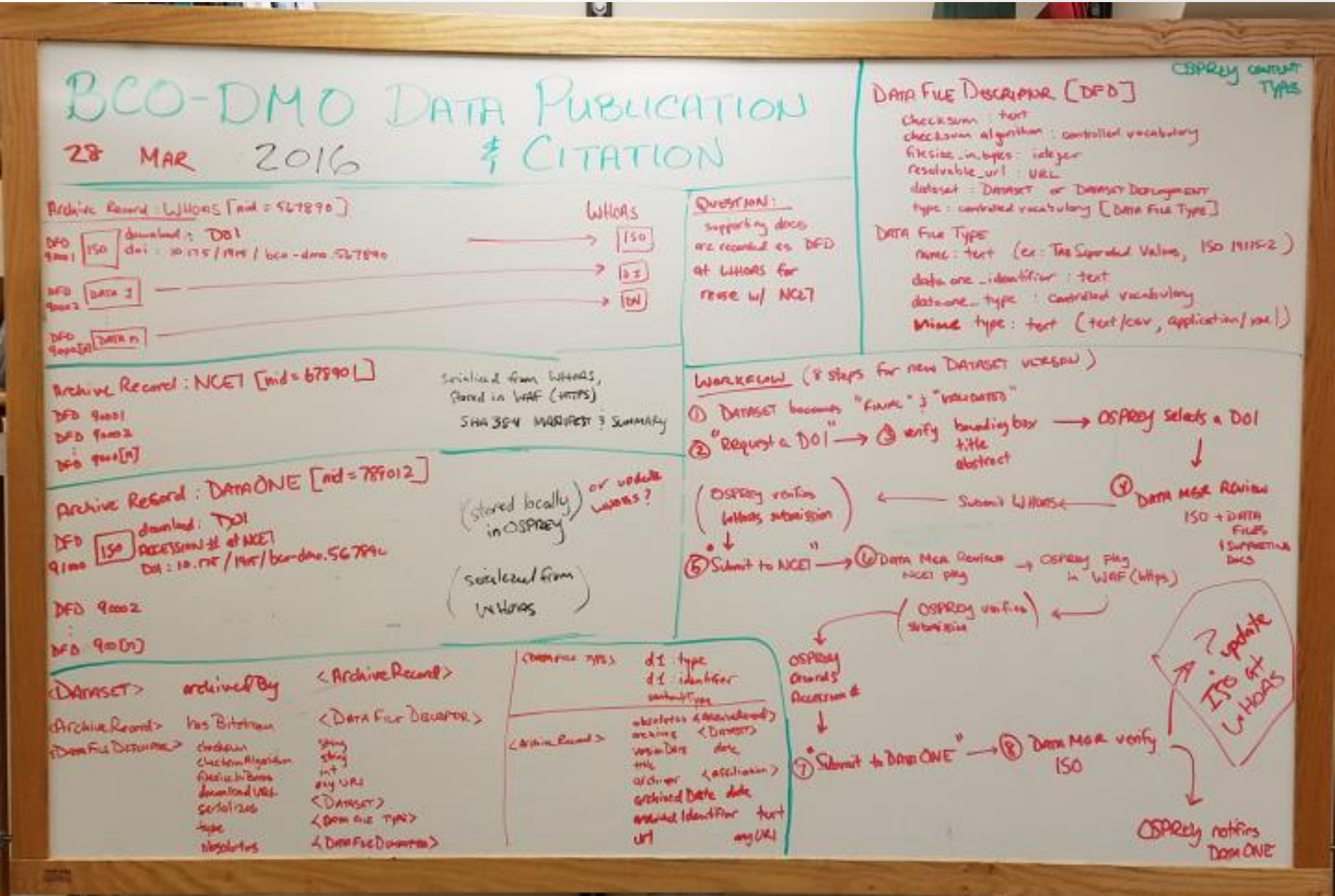
▼ Archival Copy

Archive Location	Accession ID	Archive Copy URL
National Oceanographic Data Center (NODC)	0112635	http://accession.nodc.noaa.gov/0112635

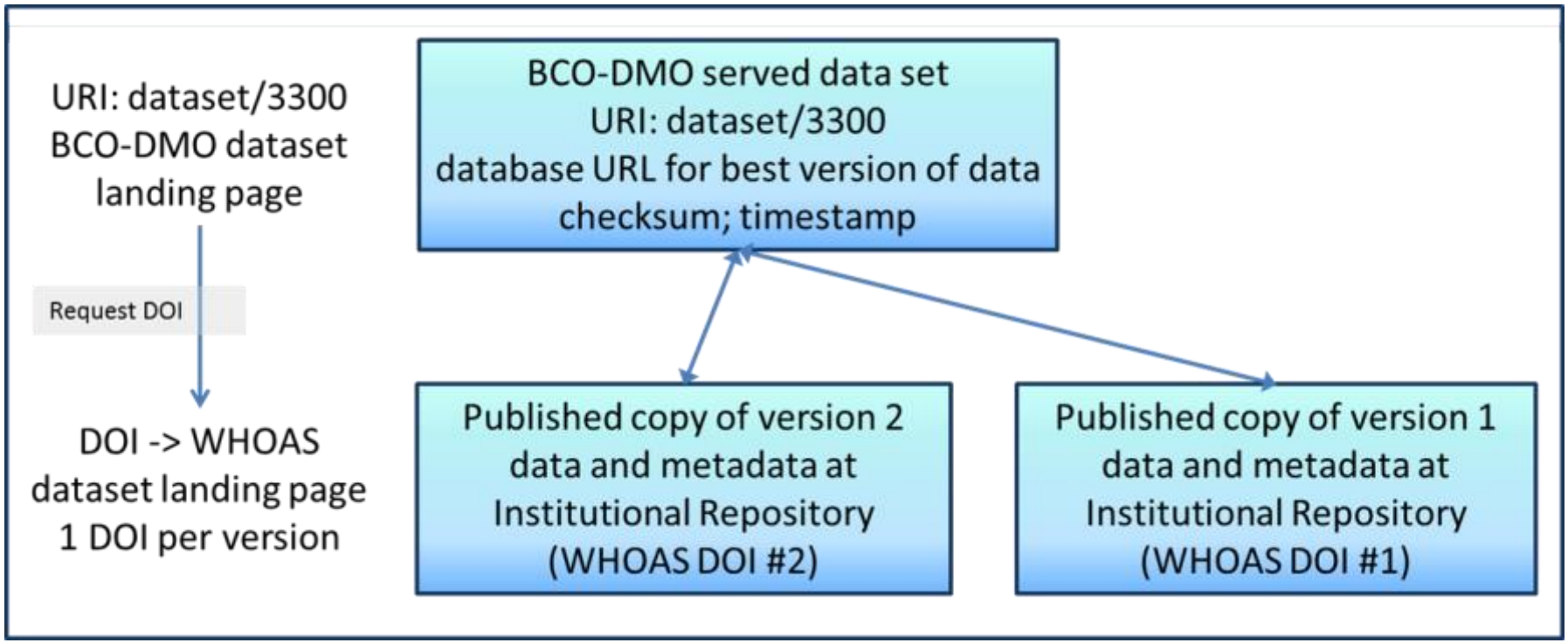
Initial Architecture Design Considerations (Jan 2016)



Modified Architecture (March 2016)

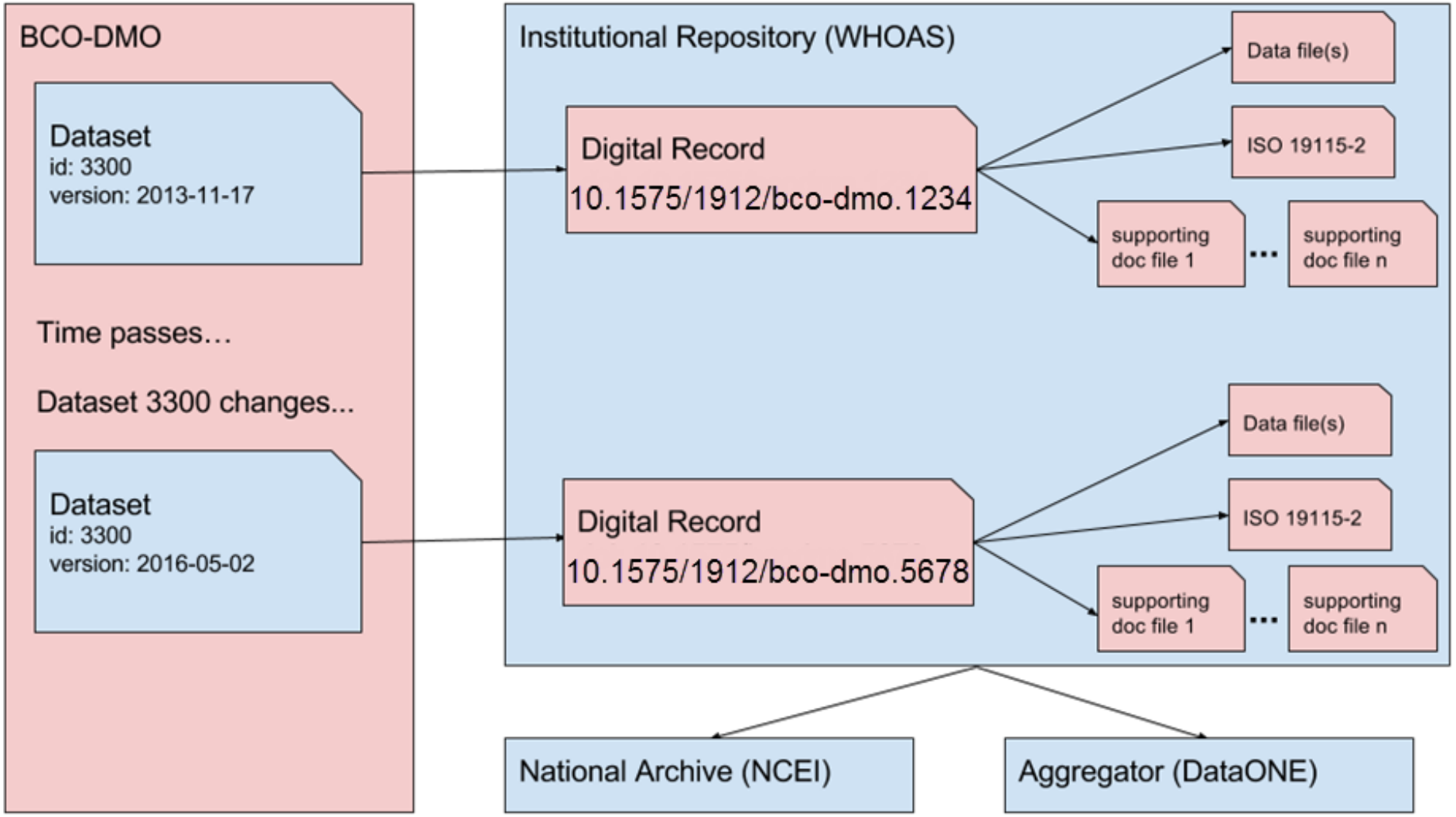


BCO-DMO Data Publication System Components

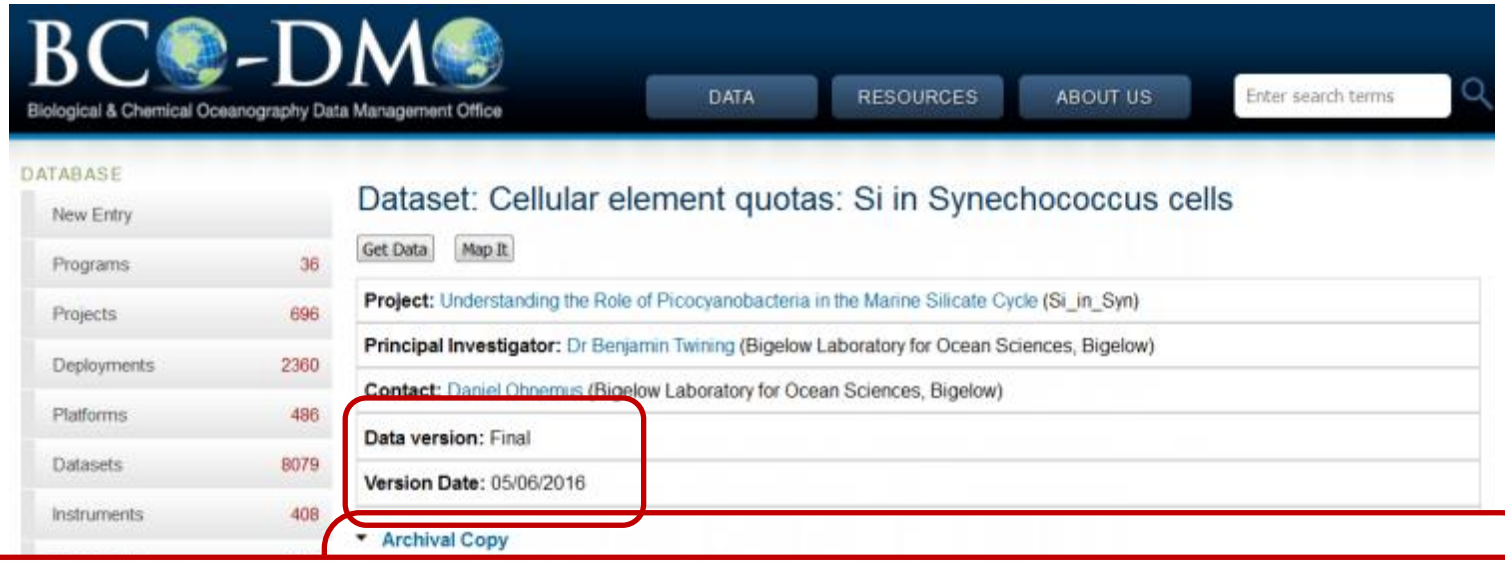


BCO-DMO publishes data to WHOAS and a DOI is assigned.
The BCO-DMO architecture now supports data versioning.

BCO-DMO Data Citation System Components



BCO-DMO Data Set Landing Page



Archival Copy

Version Date	Archive	Persistent Identifier	Date Assigned
2016-05-06	Marine Biological Laboratory/Woods Hole Oceanographic Institution Library (MBLWHOI DLA)	10.1575/1912/bco-dmo.651474	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

More Information about this dataset

Search

Search WHOAS
This Collection

BROWSE

All of WHOAS

Communities & Collections

By Issue Date

Authors

Titles

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

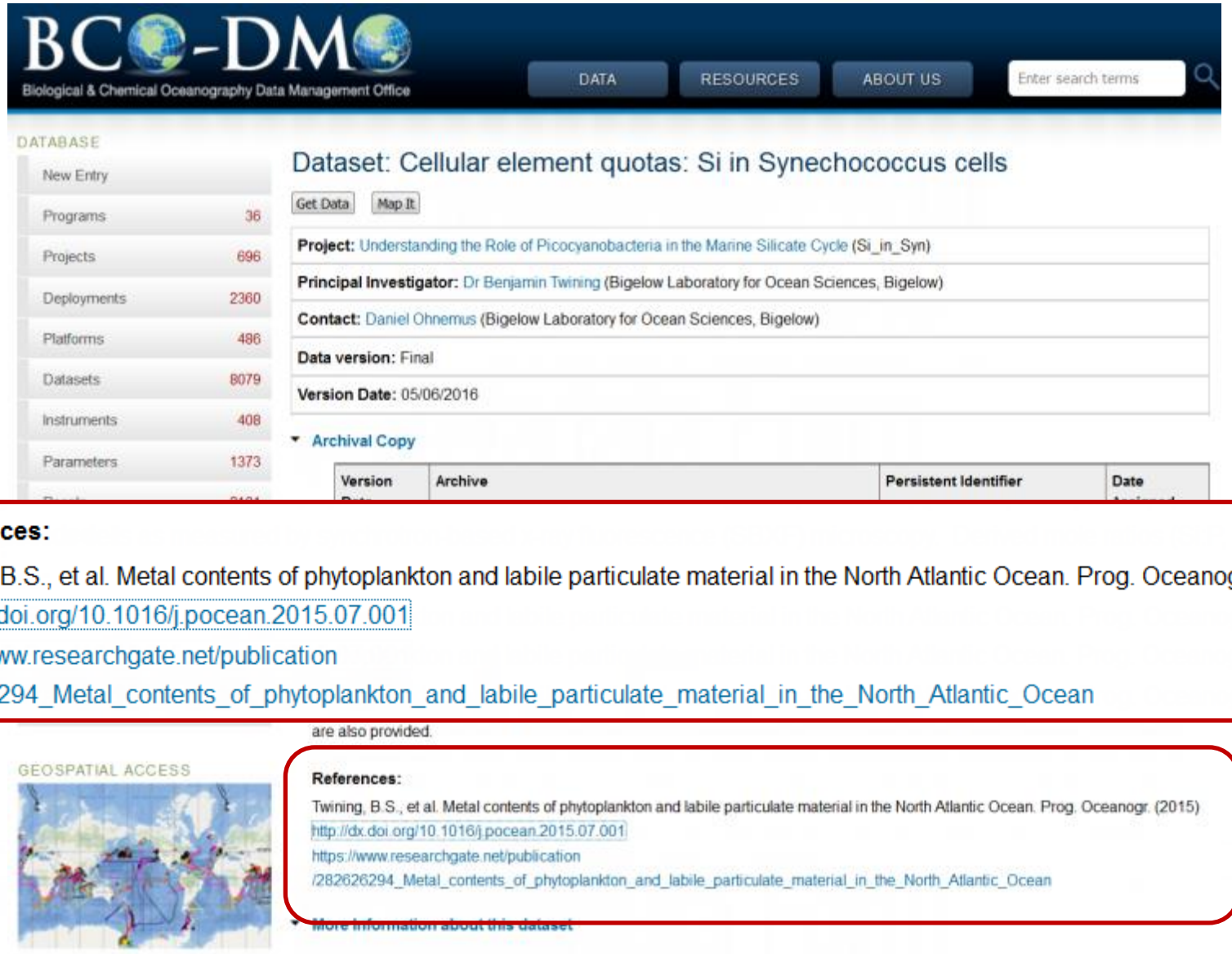
dc.description.sponsorship	NSF Ocean Sciences (NSF OCE) OCE-1131139, NSF Ocean Sciences (NSF OCE) OCE-1335012, NSF Ocean Sciences (NSF OCE) OCE-1131046
dc.publisher	Biological and Chemical Oceanography Data Management Office (BCO-DMO). Contact: bco-dmo-data@whoi.edu
dc.relation	http://lod.bco-dmo.org/id/dataset/644840

31.6691

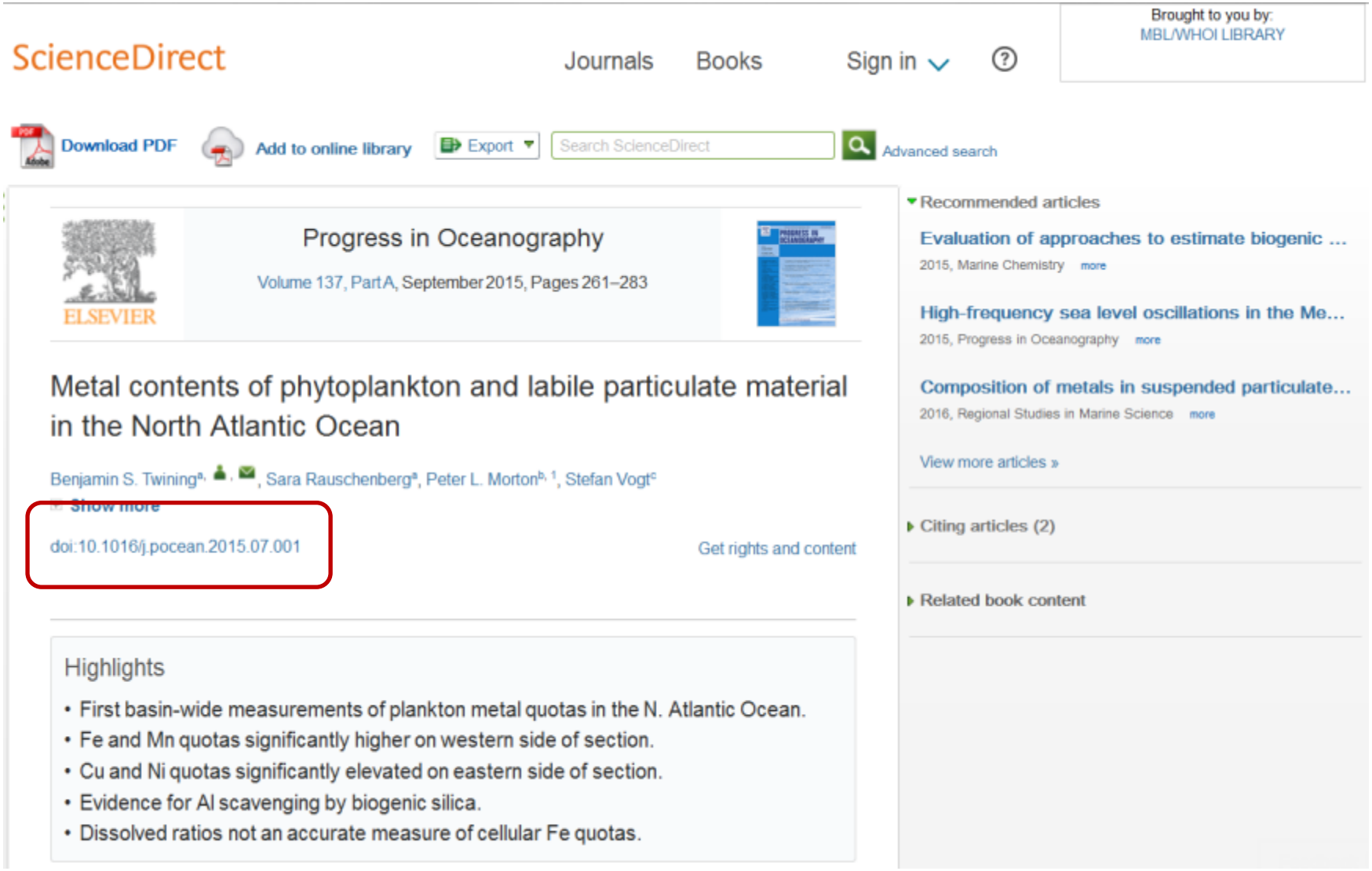
DOI
10.1575/1912/bco-dmo.651474

Keyword
SXRF; Synchrotron radiation X-Ray Fluorescence; Synechococcus; sil

BCO-DMO Data Set Landing Page



Linked to Publication via DOI



ScienceDirect Journals Books Sign in

Brought to you by: MBL/WHOI LIBRARY

Download PDF Add to online library Export Search ScienceDirect Advanced search

ELSEVIER

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[a], Sara Rauschenberg[a], Peter L. Morton[b, 1], Stefan Vogt[c]

Show more

doi: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

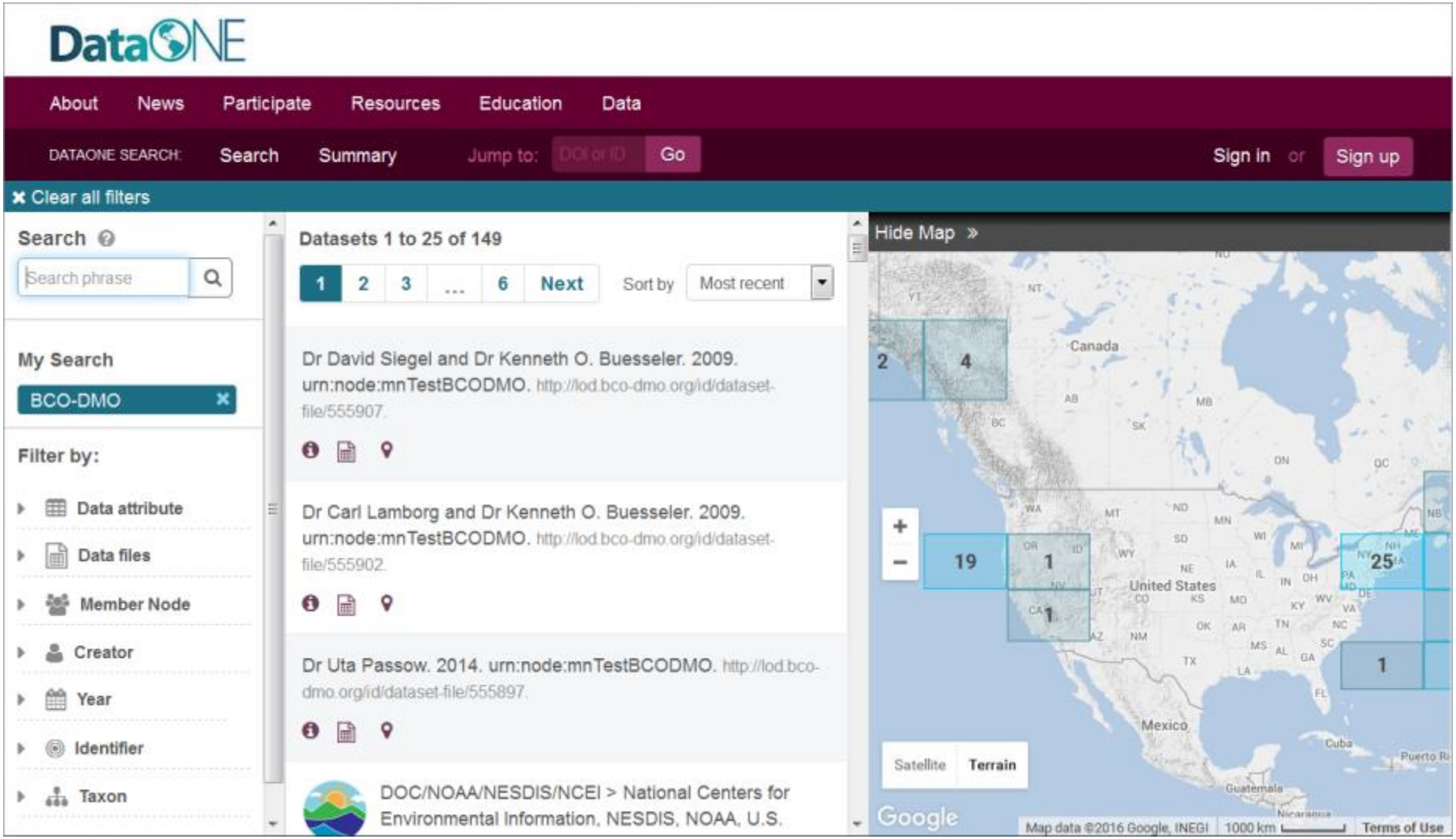
View more articles »

Citing articles (2)

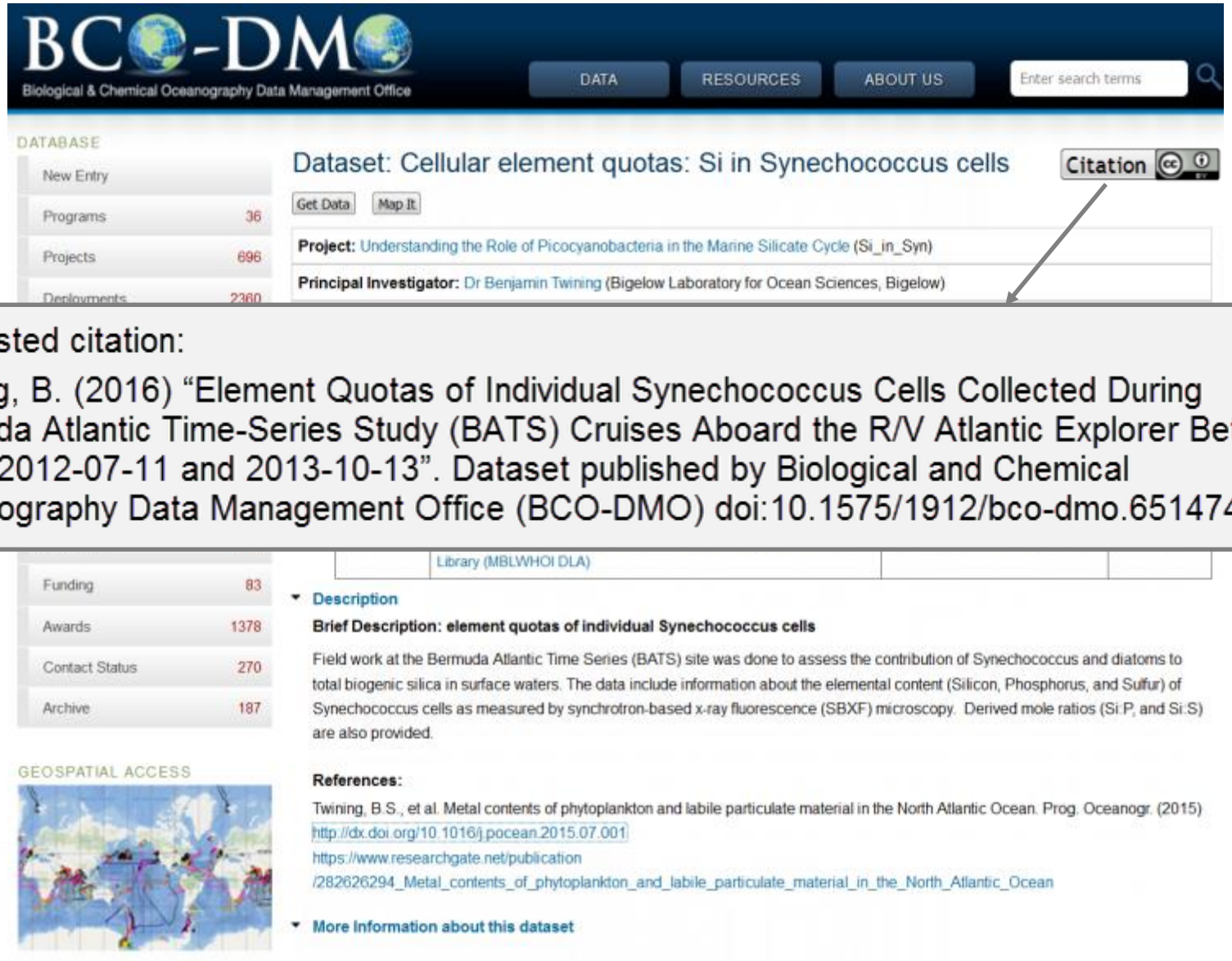
Related book content

New Capabilities … BCO-DMO becoming a DataONE Member Node

https://search.dataone.org/



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



Thank you …

- 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 @bcodmo
ORCID: 0000-0003-2129-1647 cchandler@whoi.edu

EXTRA SLIDES

- Removed these to reduce talk to 15 minutes

Adoption of Data Citation Outputs

- **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

REFERENCES

- 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

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

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

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.