

Improved Discovery and Re-Use of Oceanographic Data through a Data Management Center

Shannon Rauch

Biological and Chemical Oceanography Data Management Office
Woods Hole Oceanographic Institution

Co-authors: Molly Allison, Robert Groman, Cynthia Chandler, Charlton Galvarino, Stephen Gegg, Danie Kinkade, Adam Shepherd, Peter Wiebe, and David Glover

AGU Fall 2013
San Francisco, CA
ED43E-03



Data Discovery and Re-Use: Two perspectives

Data Contributor

Scientists, researchers, technicians, analysts who want to **share their data**.

Challenges:

- What formats and standards to use?
- How much metadata to provide?
- How will others find the data?

Data User

Scientists, researchers, technicians, analysts, students, curious individuals who want to **find data**.

Challenges:

- Where to find the data?
- How to determine if the data are appropriate for use?
- How to obtain data in a format they can use?

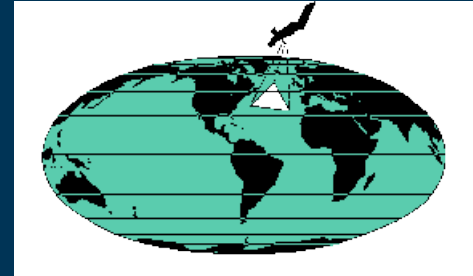
One Solution: Data Assembly Centers

- Data Assembly Center (DAC) = a facility or group of data managers that provide data management services to some community of scientists.
- Data are usually managed on short to intermediate time frames. DACs are not permanent archives.
- Services DACs might provide:
 - Collect data from different researchers on a project or in a field.
 - Assemble metadata.
 - Perform some level of QA/QC
 - Produce data products.
 - Make data accessible via the web.
 - Submit data to a national archive.

BCO-DMO Background



U.S. GLOBEC (United States **G**lobal
Ocean **E**cosystem Dynamics) DMO



U.S. JGOFS (United States **J**oint
Global **O**cean **F**lux Study) DMO



BCO-DMO

Biological & Chemical Oceanography Data Management Office

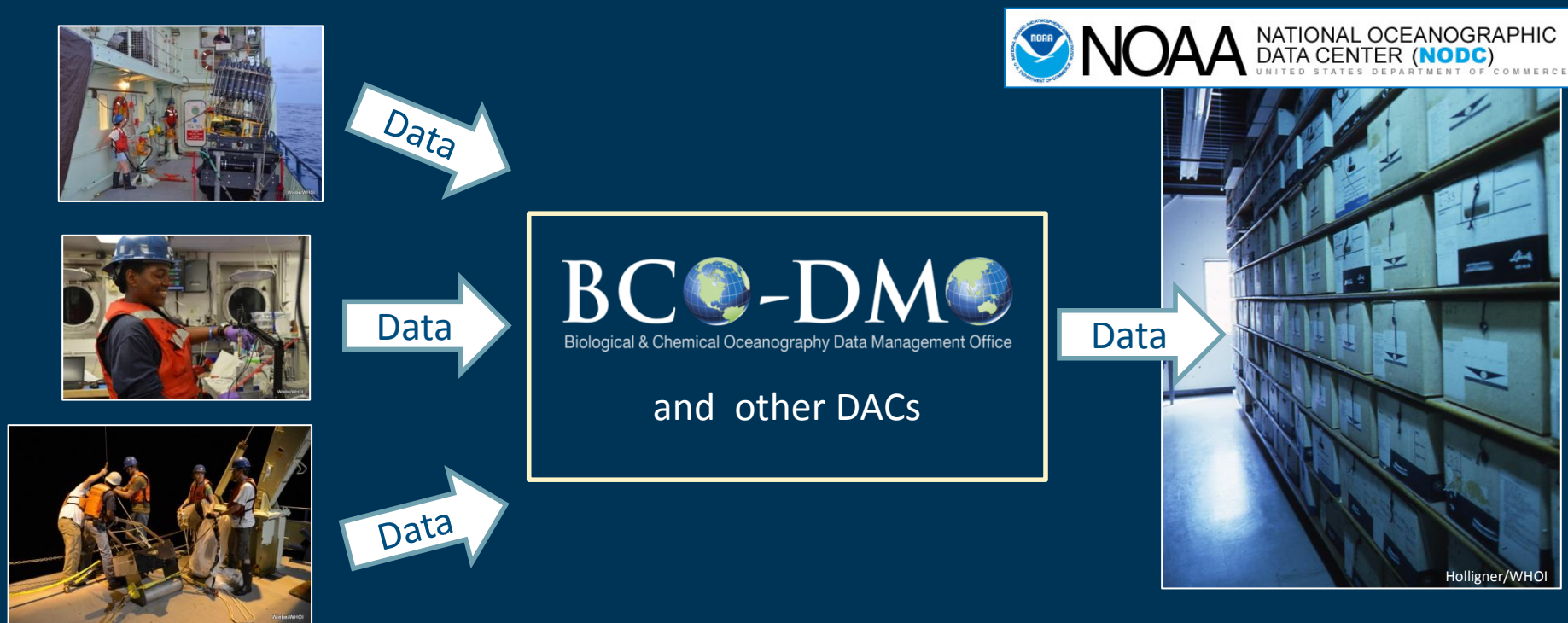
BCO-DMO provides data management, at no added cost to the investigators, to projects funded by the NSF OCE-BIO, OCE-CHEM, and PLR Antarctic Organisms & Ecosystems Programs.

Traditional Flows of Data

Data collected & analyzed,
then archived (or not).



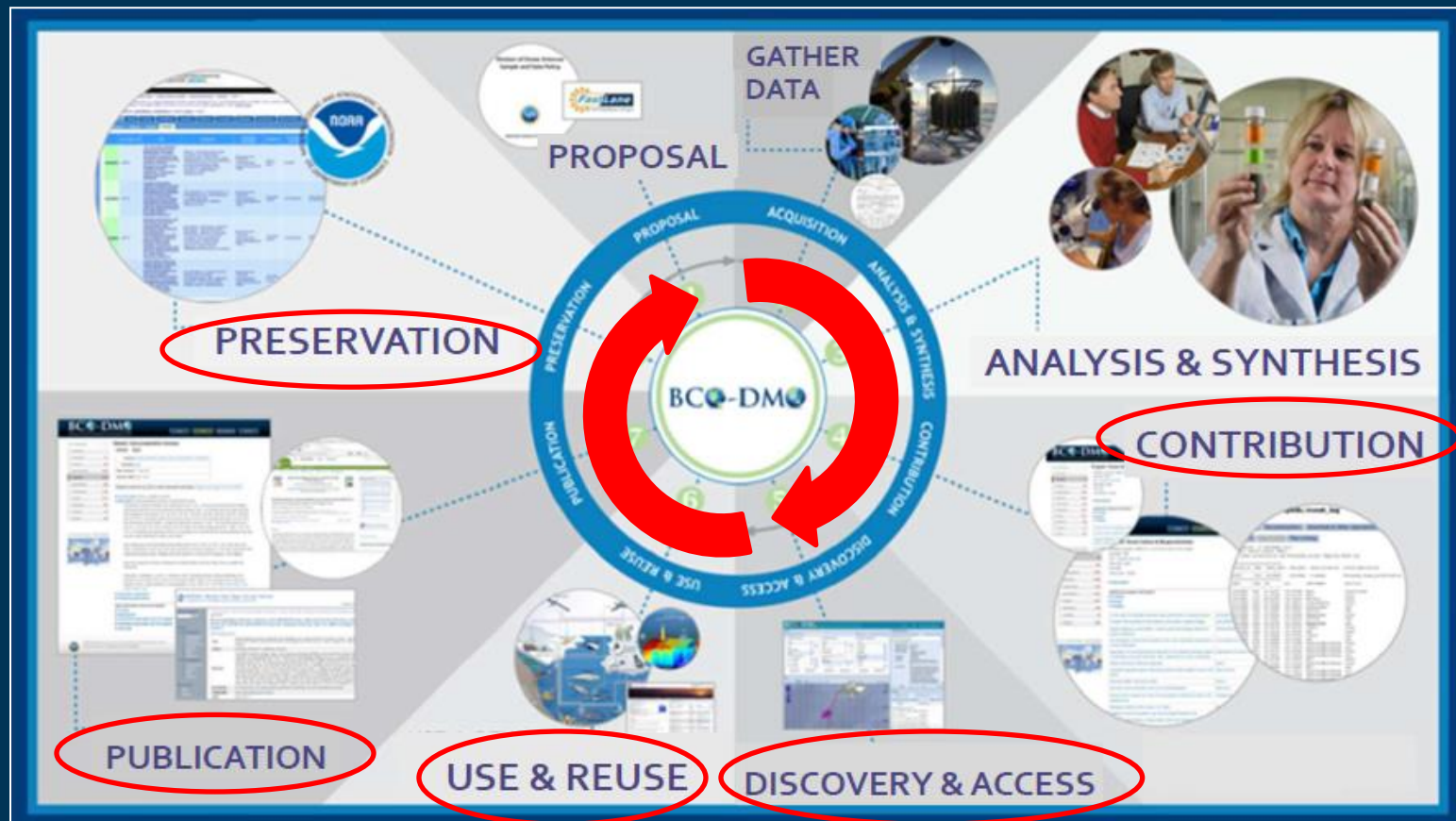
Improved Flow of Data



Data flows from scientists right to the DAC. The DAC ensures data are properly documented and archived.

BCO-DMO plays a role throughout the data lifecycle

An overarching goal is to enhance data discovery and re-use.

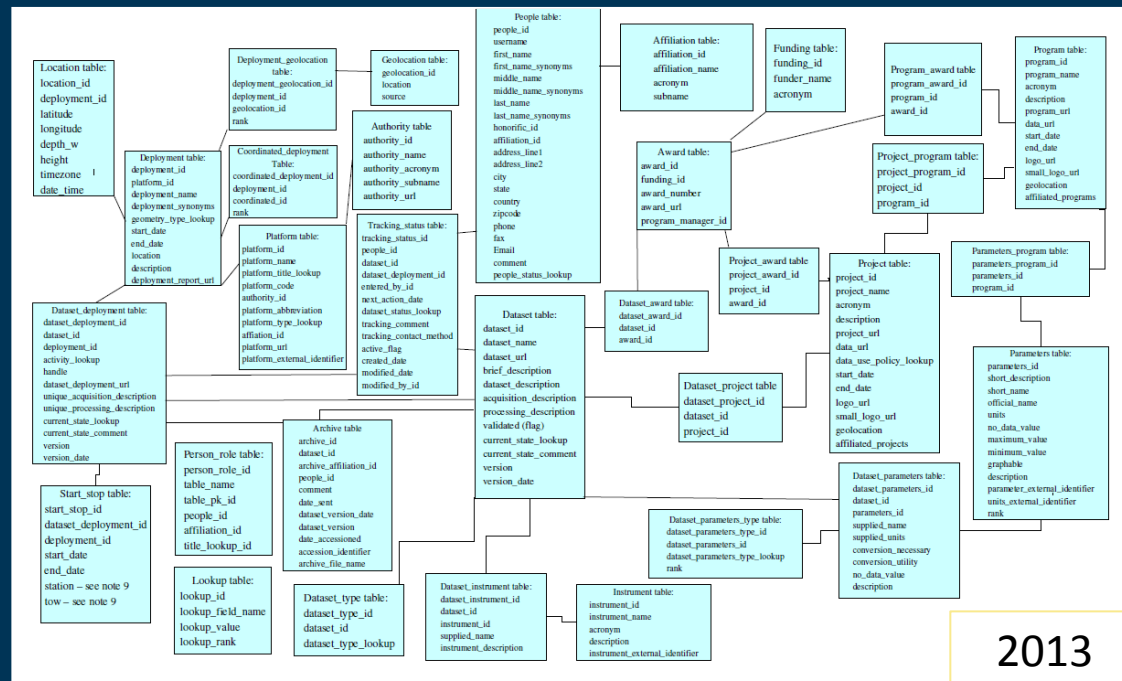


Chandler et al., AGU 2012.

Data Contribution at BCO-DMO

Data files are accepted in any format (comma- or tab-separated ASCII, Excel, MatLab, etc. as well as images and video).

Metadata, metadata, metadata – *essential* for re-use.



Controlled Vocabularies

Parameter Mapping Examples

Original data sets	BCO-DMO	SeaDataNet http://vocab.nerc.ac.uk/collection
T090C CTD T (°C)	temp	Temperature (ITS-90) of the water body /P01/current/TEMP901/
Sal00 CTD Salt	sal	Practical salinity of the water body by CTD and computation using UNESCO 1983 algorithm /P01/current/PSALST01/
DepSM Actual Depth (m)	depth	Depth below surface of the water body by profiling pressure sensor and converted to seawater depth using UNESCO algorithm /P01/current/DEPHPR01/

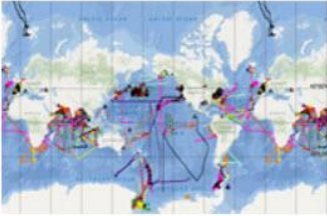
local  global

Data Discovery: text-based

BCO-DMO system uses open-access software including: Drupal, MapServer, OpenLayers, ExtJS, and MySQL.

DATABASE

Programs	26
Projects	270
Deployments	1766
Datasets	6594
Instruments	329
Parameters	1286
People	1384
Affiliations	353
Funding	54
Awards	643

GEOSPATIAL ACCESS

CONTRIBUTE DATA

Getting started

- » How-to Guide
- » FAQs

Datasets

Search

Contains all words ▼

dissolved iron

Searched for 'Contains all words'

Dataset ▲	Project
AEROSOL	Global O Aluminu (CLIVAR
dissolved_iron	Europea
Fe Speciation	Cobalt, I (CoFeM
Fe_Dissolved	Rising cl grazing (
Fe_Dissolved	Controls
GT10-11 - dFe and dFe(II)	U.S. GE
GT10-11 - Fe Mn Zn Cd and Cu	U.S. GE
GT10-11 - Fe Speciation	U.S. GE
<u>nutrients and metals</u>	Cobalt, Iron and Micro-organisms from the Upwelling zone to the Gyre (CoFeMUG)
Particulate Trace	Global Ocean Survey of Dissolved Iron and Aluminum and Aerosol Iron and

Dataset: nutrients and metals

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

Project:	Cobalt, Iron and Micro-organisms from the Upwelling zone to the Gyre
Principal Investigator:	Mak Saito (Woods Hole Oceanographic Institution)
BCO-DMO Data Manager:	Shannon Rauch (Woods Hole Oceanographic Institution)
Validated:	yes
Data version:	7 October 2009
Version date:	Mar 20, 2013

Dataset handle ID (e.g. DOI or other persistent identifier): not yet assigned

► **Current state:** Final no updates expected

▼ **Description:** Analysis of nutrients, Cobalt, Iron, and Manganese from TM bottle samples.
Analysis of nutrients, Cobalt (total and labile), dissolved Iron, and dissolved Manganese of water samples drawn from Trace Metal Rosette (TMR) bottle casts.

Related Publications:

Noble, A.E., C. H. Lamborg, D. C. Ohnemus, P. J. Lam, T. J. Goepfert, C. I. Measures, C. H. Frame, K. L. Casciotti, G. R. DiTullio, J. Jennings, M. A. Saito. 2012. Basin-scale inputs of cobalt, iron, and manganese from the Benguela-Angola front to the South Atlantic Ocean. *Limnology and Oceanography*, 57(4) 989-1010. doi:10.4319/lno.2012.57.4.0989

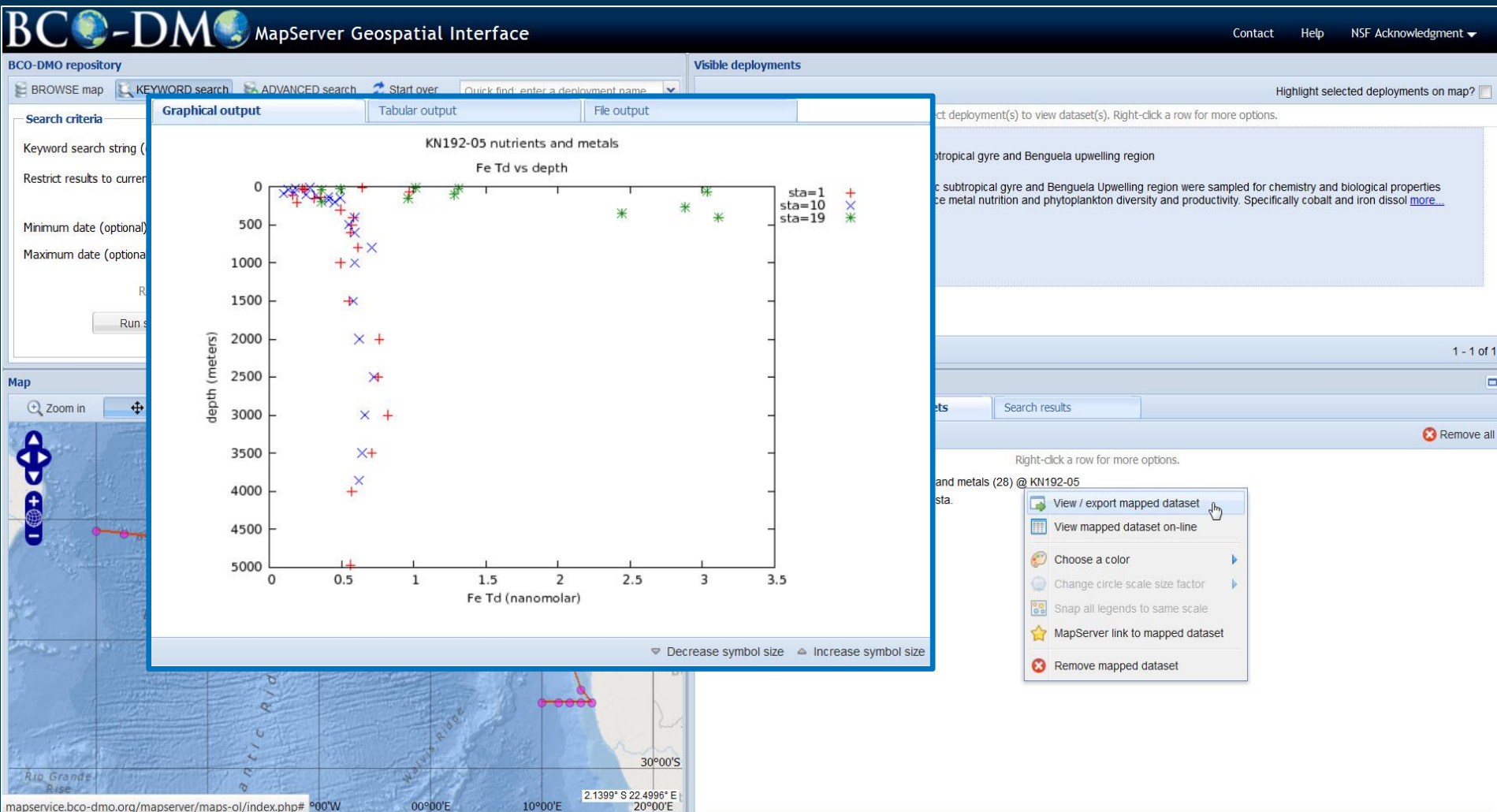
Sohm, J. A., J. A. Hilton, A. E. Noble, J. P. Zehr, M. A. Saito, and E. A. Webb. 2011. Nitrogen fixation in the South Atlantic Gyre and the Benguela Upwelling System. *Geophys. Res. Letters*. 38: L16608, doi:10.1029/2011GL048315

► **Acquisition description:**

► **Processing description:**

Dataset	Project	Validated	Data version	Version date
nutrients and metals	Cobalt, Iron and Micro-organisms from the Upwelling zone to the Gyre (CoFeMUG)	Yes	7 October 2009	
Particulate Trace	Global Ocean Survey of Dissolved Iron and Aluminum and Aerosol Iron and	No		17 August

Data Discovery: geospatial-based



Data Access, Publication, & Preservation

- Data are stored and made available for download in non-proprietary formats.
- Digital Object Identifiers (DOIs) are assigned to datasets through the WHOI-MBL Library, giving the dataset a unique persistent identifier that is citeable in publications.
- BCO-DMO ultimately deposits data at the national archive.



Summary

- A DAC, like BCO-DMO, can **facilitate data discovery** by providing the infrastructure, tools, and expertise to manage and share data, shifting some of this burden off of individual scientists.
- **Re-use of data by others is enhanced** by an effective discovery systems, open access to data in non-proprietary formats, and thorough metadata.



Happy data contributors and users!

Thank you!

www.bco-dmo.org

Email: info@bco-dmo.org



Other BCO-DMO presentations at Fall AGU:

Posters: PA31A-1814, IN31C-1508

Talks: IN31D-03, **IN52B-05 (tomorrow)**