Technologies for a FAIRer Use of Ocean Best Practices

Pier Luigi Buttigieg, Pauline Simpson,
Jay Pearlman, Peter Pissierssens, Scott Caltagirone, Mark Bushnell, Juliet Hermes, Emma Heslop, Johannes Karstensen, Frank Muller-Karger, Cristian Muñoz and Francoise Pearlman
EU Horizon 2020 project
A large scale EU Horizon 2020 research and innovation project.

62 Partners, 18 Countries
International integration of Atlantic ocean observing activities – further supporters / associates are welcome.
The main goal:
Transition from a loosely-coordinated set of existing ocean observing activities to a sustainable, efficient, and fit-for-purpose Integrated Atlantic Ocean Observing System (IAOOS)
WP6 Cross-cutting issues and emerging networks

[...] Coordinate engagement of international expertise and dissemination of best practice to improve efficiency and impact

[...] to include international expertise in the design and development, operation and maintenance of IAOOS and to disseminate this best practice
Best Practice

A documented procedure which, through experience and research, has consistently shown results superior to those achieved by other means and can be used as a benchmark, particularly if advocacy can lead to it being widely adopted

...modulo cost, feasibility, and context
...and which has a process for evolution

BPs should be catalysts for “Next Practices”
OceanBestPractices System: a global resource to facilitate harmonizing practices in ocean observation, data and information

Peter Pissierssens, Cristian Munoz, Johannes Karstensen, Jay Pearlman, Pauline Simpson, Mark Bushnell, Juliet Hermes, Emma Heslop, Frank Muller-Karger and Francoise Pearlman
The status quo...
Most potential BPs are scattered, have varying degrees of accessibility, and varying digital lifetimes.
The Mission: solutions for a FAIRer future for Ocean Best Practices

**FAIR:** Findable, Accessible, Interoperable, Reusable

Wilkinson et al. (2016) *Scientific Data* 3, DOI:10.1038/sdata.2016.18
The basis...
Disclaimer: IODE/IOC does not warrant that the information, documents and materials contained in the OceanBestPractices repository website is complete and correct and shall not be liable whatsoever for any damages incurred as a result of its use.
Findable
Accessible
Interoperable
Reusable

Wilkinson et al. (2016) *Scientific Data* 3, DOI:10.1038/sdata.2016.18
Adding semantics...
TO BE INTEROPERABLE:

1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

2. (meta)data use vocabularies that follow FAIR principles.

3. (meta)data include qualified references to other (meta)data.
Semantic tech
A rough illustration of the semantic gradient

Aiming for a high degree of expressivity, thus advanced querying and data mobilisation options

Aiming from McCreary D (2006)

Patterns of Semantic Integration. CC 2.5
UN Decade of Ocean Science for Sustainable Development 2021-2030

“To gather ocean stakeholders worldwide behind a common framework that will ensure ocean science can fully support countries in the achievement of the Sustainable Development Goal 14 on the ocean.”

Sustainable Development Goals Interface Ontology

Building data systems to support interfaces to the Global Goals, targets, and indicators

#OceanScienceDecade
A fragment of envoPolar’s nivation semantics

Each node links to many others in ENVO and other interoperating ontologies
Welcome to the CMECS Catalog of Units!

Use this database to browse the CMECS classification and to get definitions for individual CMECS Units. This database contains the units that were published in the Coastal and Marine Ecological Classification Standard.

Search/Browse classification
About the classification
Download the CMECS Standard
Download the Data:

Excel
Access 2007

NEW: Propose update to classification
NEW: Search for proposals
Testing ontological expression of CMECS, clarifying ambiguities, and creating development cycle with the CMECS teams

CMECS

https://cmeccatalog.org/cmeccs/classification/unit/666.html

semi-diurnal tidal flow
OceanBestPractices
Repository of Ocean Community Practices in
Ocean Research, Observation and
Data/Information Management

BROWSE
All of OceanBestPractices
Communities & Collections
By Issue Date
Authors
Titles
Subjects

OBO Foundry Ontologies
ENVO, CHEBI, SDGIO

Their BP
Author  Author
Part 1

My BP
Author  Author
Part 1

BODC sensor vocabs
Rapidly access document parts linked to terms of interest
9.1 Platform Description

Grabs and box corers both use receptacles to collect sediment after they are dropped to the seafloor. While the scooping motion of grabs disrupts unconsolidated sediment to various degrees, box corers return largely undisturbed samples of the sediment strata (Eleftheriou and McIntyre 2005). Grabs and box corers target surface sediment and associated porewater and fauna. They are typically deployed over sandy or muddy substrates, although some grabs can collect gravel or cobbles.

There is no single type of grab or box corer suitable for all environments, and selection of the most suitable type depends on the biological or physical target, substrate, depth, and vessel capabilities (Narayanaswamy et al. 2016). Acquired data can be quantitative (e.g. volumetric or mass specific) or semi-quantitative due to inconsistencies in sample volume and sediment disruption due to bow waves or other gear effects (Blomqvist 1991). For these reasons, this manual does not mandate specific gear types. There are numerous references to help facilitate decisions regarding grab and box corer equipment for a given marine survey (Riddle 1989, Eleftheriou and Moore 2005, Danovaro 2010, Narayanaswamy et al. 2016). Nevertheless, for monitoring purposes, it is preferable to maintain consistent gear through time and space, and we therefore recommend this where possible.

9.2 Scope

This Grab and Box Corer Field Manual includes gear designed to sample unconsolidated sediment and organisms on the seafloor. General steps are outlined in Figure 9.1 and described in detail in the sections below.

The samples collected by grabs and box corers can be used to derive a range of physical, chemical, and biological parameters (Eleftheriou and McIntyre 2005), and each of these parameters requires a
Explore knowledge neighbourhood to enhance discovery.
Leveraging contextual domain knowledge to enhance search and discovery

Constrain / Expand searches assisted by knowledge models
Ensuring fitness-for-purpose...
Ontologies / Vocabularies

Merged triplestore

Tagging Module

Title
Author

Title
Author

Part 1

Semantic + ElasticSearch interface

Link back to original documents

Raw text extraction

Submission interface

Ocean Best Practices

Full BP

BP Developers

Ocean Observation Section for BPs

Ontologies / Vocabularies
Ocean Observation Section for BPs

BP Developers

Raw text & Metadata extraction

NLP module
- N-grams
- Unstructured keywords
- New classification BPs

Submission interface

Link back to and microcredit original documents

Feedback / co-development
- Ontologies / Vocabularies

OceanBestPractices

frontiers in Marine Science

Ocean Observation Section for BPs
Concluding remarks

We’re well on the way to a semantically enhanced, international BP repository for the Ocean Observation community

The near future...

• Incorporate more semantic and terminology resources to improve coverage, focusing on interoperation and relevance to ocean observing
• Refine NLP and tagging routines for improved accuracy and knowledge harvesting
• Provide community guidelines on submitting parseable best practices
• Use semantic tags to link BPs to oceanographic data holdings and streams
• Improve user experience of the OBP system to simplify interaction with ontologies, while preserving their expressivity
Opening up BPs to informatics and linking them to data and information holdings

- Retrieve the common **sensors** across the protocols “McMussel mussel bed survey” and “OysterTech oyster reef survey”
- Retrieve all BPs that use the **software** “ArcGIS”
- Retrieve the **advantages and limitations** of the protocol titled “Sampling seawater for marine microbes”
- Retrieve the protocols and the list of **reagents** for documents authored by E Hemmingway

Inspired by [https://smartprotocols.github.io/](https://smartprotocols.github.io/)
Acknowledgements

Antje Boetius
Felix Janssen

Jacqueline McGlade
Ludgarde Coppens
Dany Ghafari

Chris Mungall (UCB)
Ramona Walls (CyVerse)
Evangelos Pafilis (HCMR)
Mark Schildhauer (NCEAS)
Barry Smith (SUNY)
Suzanna Lewis (UCB)
and many, many contributors

Best Practices Working Group

Mark Bushnell
Juliet Hermes
Emma Heslop
Peter Pissiersens
Cristian Muñoz
Francoise Pearlman
Jay Pearlman
Pauline Simpson
Frank Muller-Karger

IOOS
SAEON/JCOMM
 UNESCO
 UNESCO
 SOCIB
 IEEE
 IEEE
 IODE
 MBON

CMECS: Mark Finkbeiner (NOAA)
Derek Sowers (NOAA)

BODC: Alexandra Kokkinaki