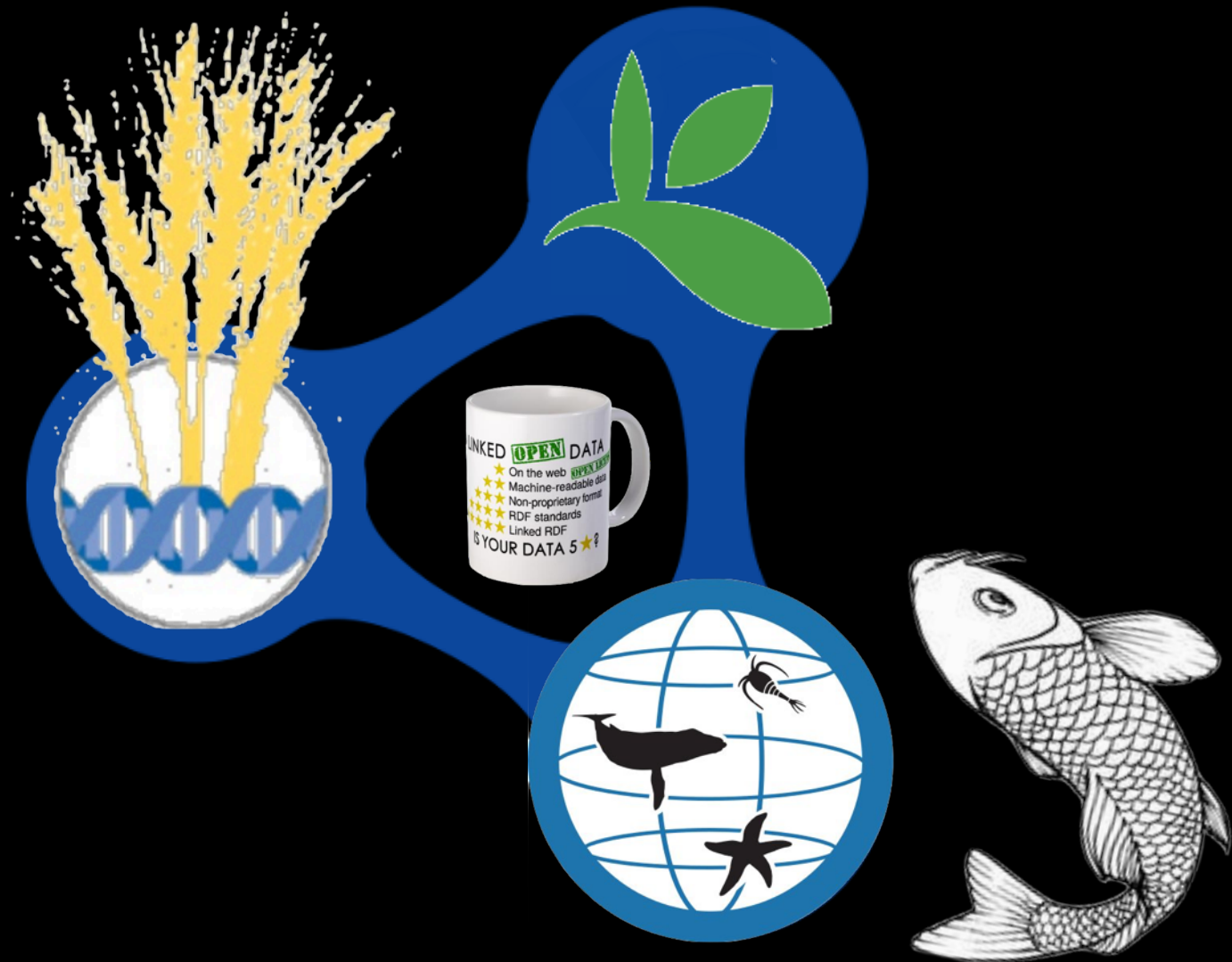


Event core for measurement or fact data

Belpo, Brussel, January 2018



Event core and measurements

- **Event core** (since 2015) opens up new functionality for data published in GBIF!
- Limitations of the **star-schema** is a frequently mentioned obstacle...
 - E.g. when connecting measurement or facts (MoF) only core entities can be described.
- "**OBIS eMoF**" provides flexible solutions that could be further generalized and suggested as the new updated "**GBIF:MoF**" standard...?

DwC archive – application schema

- **DwC Event (since 2015)**
 - http://rs.gbif.org/core/dwc_event_2016_06_21.xml
- DwC Occurrence
 - http://rs.gbif.org/core/dwc_occurrence_2015-07-02.xml
- DwC Measurement or fact
 - http://rs.gbif.org/extension/dwc/measurements_or_facts.xml
- **OBIS Extended measurement or fact**
 - http://rs.gbif.org/extension/obis/extended_measurement_or_fact.xml



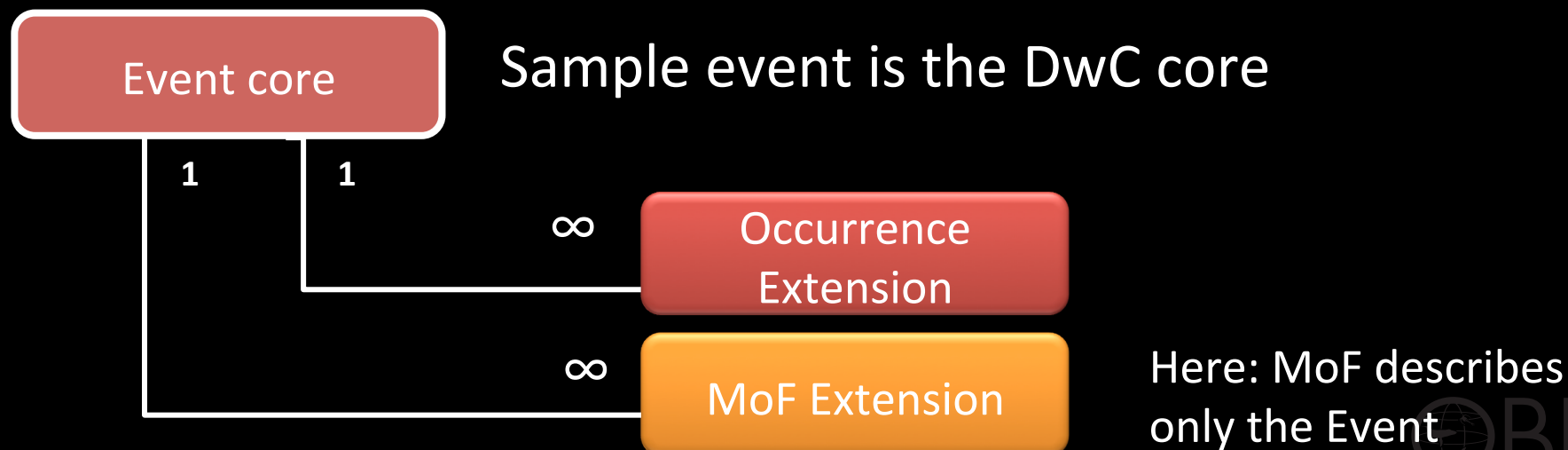
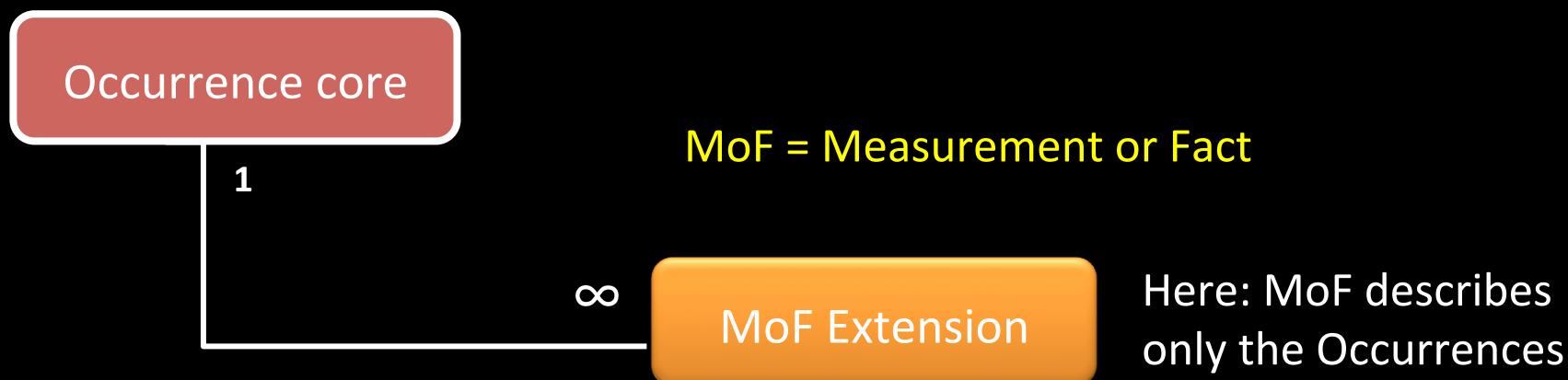
Use case : Measurements for both events and occurrences in the same dataset

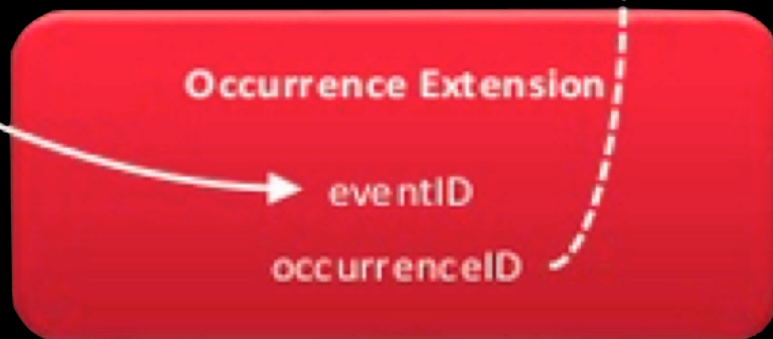
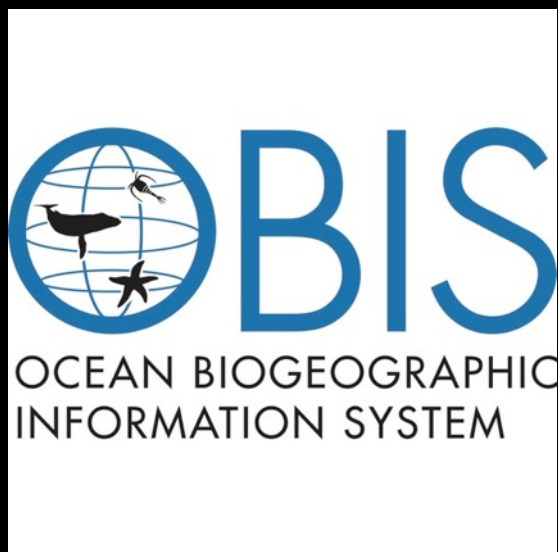
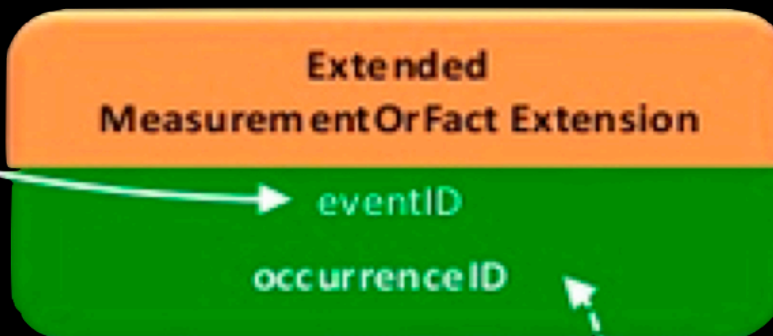
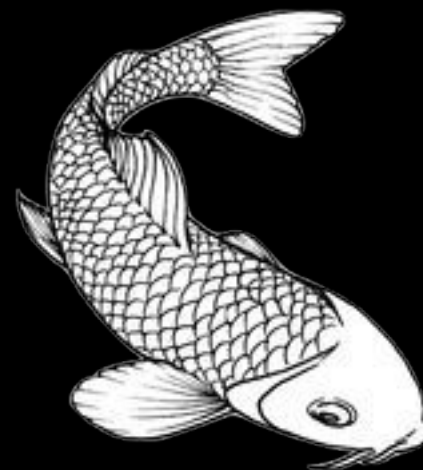
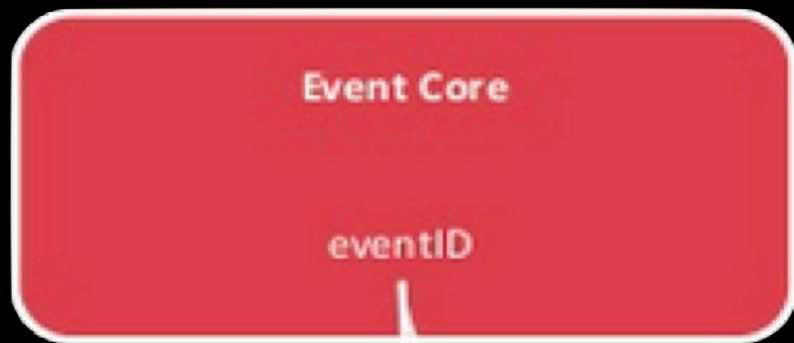
OBIS expeditions are sampling more data than only organism occurrences (e.g. environment).

--> development of a new extended measurement and fact extension (for DwC-archives)



Star Schema (DwC-archive)







Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences

▼ Daphnis De Pooter, Ward Appeltans, Nicolas Bailly, Sky Bristol, Klaas Deneudt, Menashè Eliezer, Ei Fujioka, Alessandra Giorgetti, Philip Goldstein, Mirtha Lewis, Marina Lipizer, Kevin Mackay, Maria Marin, Gwenaëlle Moncoiffé, Stamatina Nikolopoulou, Pieter Provoost, Shannon Rauch, Andres Roubicek, Carlos Torres, Anton van de Putte, Leen Vandepitte, Bart Vanhoorne, Matteo Vinci, Nina Wambiji, David Watts, Eduardo Klein Salas, Francisco Hernandez

The extended MeasurementOrFact extension allows for parameter standardization (linking to external URI of a controlled vocabulary)

ID: Link to the core record (here: a link to the event)

occurrenceID (new): identifier to link the eMoF with the occurrence extension

measurementType: The nature of the measurement or fact (MoF)

measurementTypeID (new): identifier for the measurementType (global unique identifier, URI)

measurementValue: the value of the MoF

measurementValueID (new): identifier for facts reported as measurementValue (URI)

measurementAccuracy: description of the potential error or accuracy of the MoF

measurementUnit: measurement unit (SI unit; meter, gram, etc)

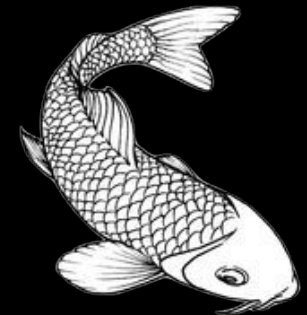
measurementUnitID (new): identifier for the measurementUnit (URI)

measurementDeterminedDate: date on which the MOF was made

measurementDeterminedBy: list of people, groups, or organizations who recorded the MoF

measurementMethod: description or reference to (publication, URI) for the method or protocol used

measurementRemarks: Comments or notes to the MoF



Introduction

References

Quick Reference Guide

Simple Darwin Core

RDF Guide

Table of Contents

Introduction

Implementation Guide

Term reference

Text Guide

XML Guide

Namespace Policy

Complete History

Decision History

Mapping to ABCD

Mapping to Old Versions

Darwin Core RDF Guide

Title: Darwin Core RDF Guide

Date Issued: 2015-03-27

Date Modified: 2015-06-15

Abstract: This guide is intended to facilitate the use of Darwin Core terms in the Resource Description Framework (RDF). It explains basic features of RDF and provides details of how to expose data in the form of RDF using Darwin Core terms and terms from other key vocabularies. It defines terms in the namespace <http://rs.tdwg.org/dwc/iri/> which are intended for use exclusively with non-literal objects.

Contributors: Steve Baskauf (TDWG RDF/OWL Task Group), John Wieczorek (TDWG Darwin Core Task Group), John Deck (Genomic Biodiversity Working Group), Campbell Webb (TDWG RDF/OWL Task Group), Paul J. Morris (Harvard University Herbaria/Museum of Comparative Zoölogy), Mark Schildhauer (National Center for Ecological Analysis and Synthesis) (KUNHM)

Legal: This document is governed by the standard legal, copyright, licensing provisions and disclaimers issued by the Taxonomic Databases Working Group.

Part of TDWG Standard: <http://www.tdwg.org/standards/450/>

Creator: Darwin Core and RDF/OWL Task Groups

Identifier: <http://rs.tdwg.org/dwc/2015-06-02/terms/guides/rdf/>

Latest Version: <http://rs.tdwg.org/dwc/terms/guides/rdf/>

Replaces: <http://rs.tdwg.org/dwc/2015-03-27/terms/guides/rdf/>

Document Status: Current Standard

dwciri: namespace

The **dwciri**: namespace is intended for use exclusively with non-literal objects.

<http://rs.tdwg.org/dwc/iri/>

Baskauf S, Wieczorek J, Deck J, Webb C, Morris PJ and Schildhauer M (2015) Darwin Core RDF Guide, Biodiversity Information Standards <http://rs.tdwg.org/dwc/terms/guides/rdf/>

Baskauf SJ, Wieczorek J, Deck J, Webb CO (2016) Lessons learned from adapting the Darwin Core vocabulary standard for use in RDF. Semantic Web 7(6): 617-627, DOI: 10.3233/SW-150199

dwciri: and dwc: terms

- If a term in the **dwciri:** namespace has an analogue in the **dwc:** namespace having the same local name, the dwciri: term will have the **same meaning** as its dwc: counterpart, but is intended to have an **IRI-reference object**.

Darwin Core RDF guide and the dwciri: namespace term already have alternatives for some of the new terms minted by the OBIS eMoF.

emof = <http://rs.iobis.org/obis/terms/>

dwciri = <http://rs.tdwg.org/dwc/iri/>

emof:measurementTypeID = dwciri:measurementType

emof:measurementUnitID = dwciri:measurementUnit

emof:measurementValueID (no analogue in dwciri - useful addition)

see also:

dwciri:measurementDeterminedBy

dwciri:measurementMethod

dwciri:samplingProtocol

DwC RDF Guide: http://rs.tdwg.org/dwc/terms/guides/rdf/index.htm#2.9_MeasurementOrFact_instances

Linking MoF instances to resources

- *"Darwin Core does not provide the object properties that would be required to describe precisely how the `dwc:MeasurementOrFact` instance was related to the `dwc:Occurrence` instance, the `dwc:Organism` whose occurrence was recorded, or the `dwc:PreservedSpecimen` that was collected".*
- Examples of the DwC RDF Guide use **dcterms:relation** to link MoF instances to the respective resources.

DwC RDF Guide MoF example 31

Example 31 - Turtle

```
<http://mczbase.mcz.harvard.edu/guid/MCZ:Mamm:67231>
  a dwc:Occurrence;
  dcterms:relation [a dwc:MeasurementOrFact ;
    dwc:measurementType "tail length" ;
    dwciri:measurementType <http://purl.obolibrary.org/obo/VT_0002758> ;
    dwc:measurementValue "25"^^xsd:int ; dwc:measurementUnit "mm" ;
    dwciri:measurementUnit <http://mimi.case.edu/ontologies/2009/1/
UnitsOntology#millimeter> ;
    dwc:measurementAccuracy "0.5"^^xsd:decimal ;
    dwc:measurementDeterminedDate "2009-08-22"^^xsd:date ;
    dwc:measurementDeterminedBy "Ryan B Stephens" ;
    dwciri:measurementDeterminedBy <http://scholar.google.com/citations?
user=RA5UdjoAAAAJ> ;
    dwc:measurementMethod "unspecified" ;
    dwciri:measurementMethod <http://purl.obolibrary.org/obo/MMO_0000160> ;
    dwc:measurementRemarks "Accuracy from significant digits."@en].
```

Unintended consequences of ID terms

- In its normative definition, each Darwin Core ID term is declared to be `rdfs:subPropertyOf dcterms:identifier`.
- In RDF, the purpose of a subproperty declaration is to allow a client with reasoning capability to infer a triple containing a broader (and presumably more well-known) property.
- (Applicable to RDF-serialization, less so to text- or XML-serialization).

DwC RDF Guide **ID term** example 23

Example 23 : Turtle (incorrect usage)

```
<http://guid.mvz.org/identifications/23459>  
  a dwc:Identification; dwc:identifiedBy "Richard Sage";  
  dwc:dateIdentified "2000"^^xsd:gYear;  
  dwc:taxonID "http://lsid.tdwg.org/  
urn:lsid:catalogueoflife.org:taxon:d79c11aa-29c1-102b-9a4a-00304854f  
820:col20120721".
```

possible incorrect inference:

```
dcterms:identifier "http://lsid.tdwg.org/  
urn:lsid:catalogueoflife.org:taxon:d79c11aa-29c1-102b-9a4a-00304854f  
820:col20120721".
```

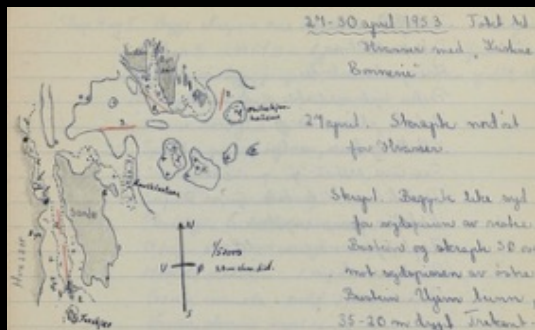

Use case : linking between multiple archives

Joint collecting expeditions collecting material that end up in different museums, herbaria, and genebanks.

- (1) **Collecting events** published in one DwC-archive
- (2) **Collection specimens** in other DwC-archive(s)
- (3) ... and *experiments* in a third set of DwC-archives

Example of a joint, collaborative collection mission: same sampling event – but specimens sent to different museum collections

Dannevig collections (280 collecting events)
<http://doi.org/10.15468/hwvr0m>



Material from the same expedition → 4 herbaria, and 4 datasets (DwC-A).

UiT : Arthropod collection
<http://doi.org/10.15468/vjkvxg>

NTNU : Algae herbarium
<https://doi.org/10.15468/elzhsu>

UiB : Cephalaspidean Gastropods
<http://doi.org/10.15468/sbomfu>

UiO : Dannevig- and Drobak-
collections Polychaeta
<https://doi.org/10.15468/y6cctp>



Proposing a new use case:

Measurement events as dwc:Event

Experiments with living specimens at a location and date that is different from the collecting event.

