

Metadata

Sarah Faulwetter & Nicolas Bailly

A Practical introduction to acquisition, validation,
quality control and access to (biodiversity) data

Vlissingen, Netherlands, 14.06.2015



What is metadata?



Data about data

data describing other data

additional information

summary of the dataset

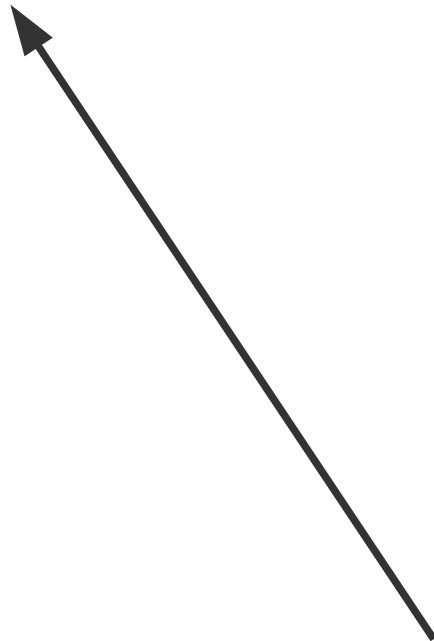
Data about data

data describing other data

additional information

summary of the dataset

Data about data



context

Types of metadata



Record level



Dataset level

Embedded



External

Depends on
parameters



Constrained by
metadata standards

Record level



Dataset level

Embedded



External

Depends on
parameters



Constrained by
metadata standards

*e.g. file size, creation date, geographic /
temporal / taxonomic scope, ...*

computed from
the dataset



added by humans

*e.g. title, description, licenses,
persons and their roles, ...*

Purpose of metadata



retrievability and discoverability



retrievability and discoverability

promote re-use of data
-> avoid duplication of efforts



retrievability and discoverability

promote re-use of data
-> avoid duplication of efforts



data interpretation



retrievability and discoverability

promote re-use of data
-> avoid duplication of efforts



data interpretation

process of information creation
can be understood and repeated



retrievability and discoverability

promote re-use of data
-> avoid duplication of efforts



data interpretation

process of information creation
can be understood and repeated

assess fitness for use



retrievability and discoverability

promote re-use of data
-> avoid duplication of efforts



data interpretation

process of information creation
can be understood and repeated

assess fitness for use



location, accessibility and terms of use



retrievability and discoverability

promote re-use of data
→ avoid duplication of efforts



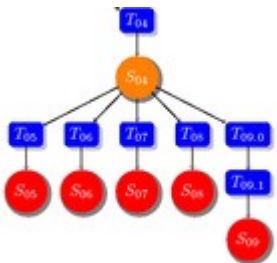
data interpretation

process of information creation
can be understood and repeated



location, accessibility and terms of use

assess fitness for use



relationship with other resources

What / when to record ?

Why?

What?

How?

Where?

When?

Who?

Why?

What?

How?

Where? + other information

When?

Who?

Why?

What?

How?

Where?

When?

Who?

+ other information



e.g. automatically recorded
metadata, licenses, related info...

Why?

What?

How?

Where?

When?

Who?

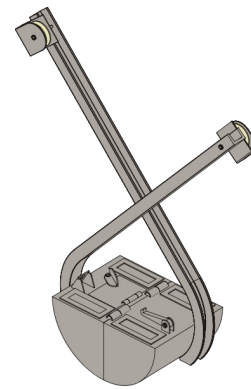
often determined by metadata
standard

+ other information

e.g. automatically recorded
metadata, licenses, related info...



general
description



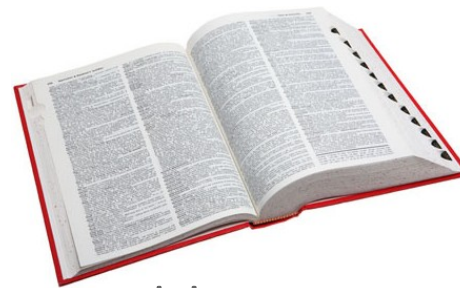
material &
methods



contributors



publications



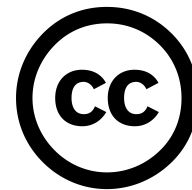
subject area
(keywords)



dataset citation



temporal, spatial [&
taxonomic] coverage



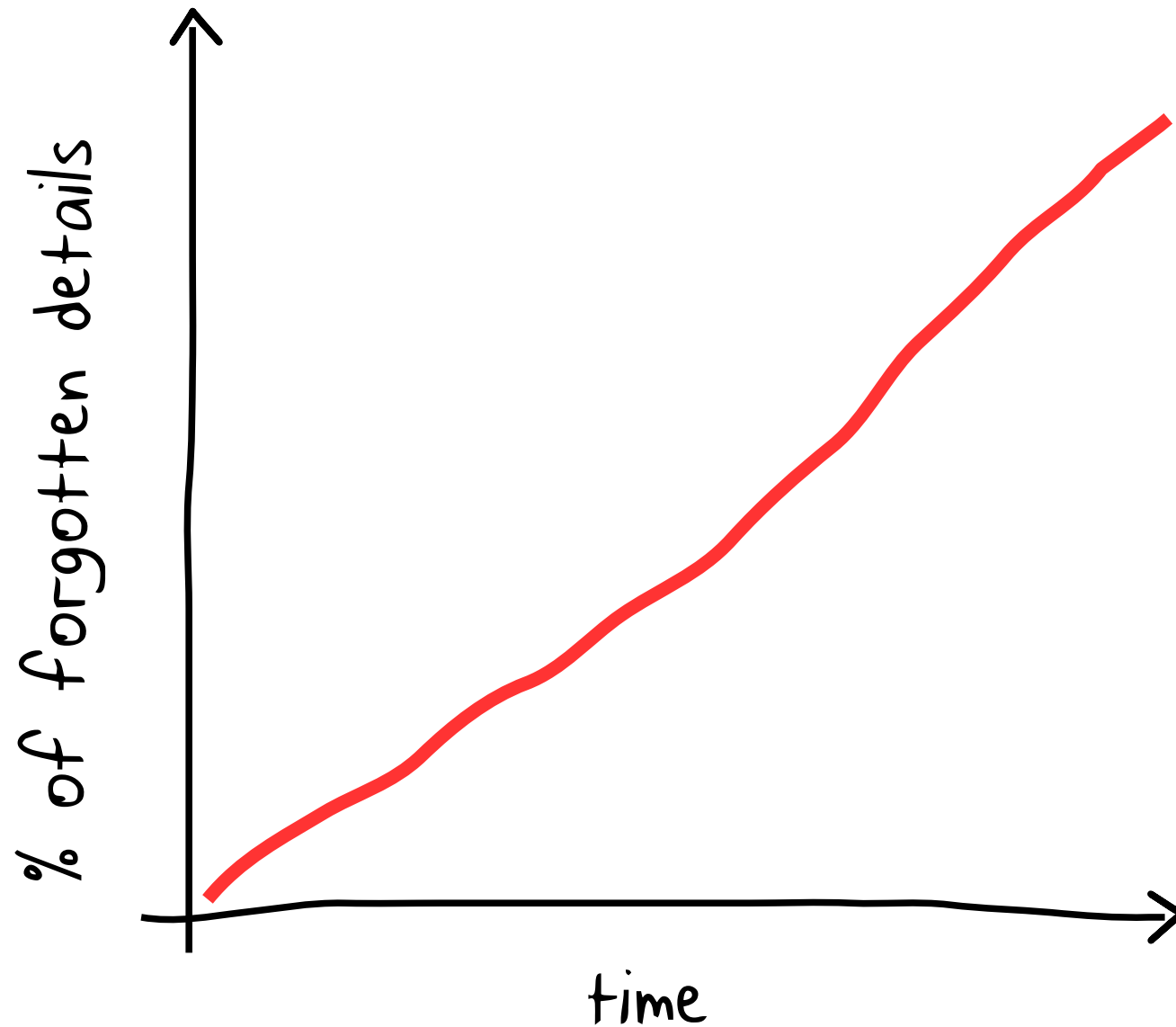
terms of use /
licenses



version

“I'll do the metadata stuff later....”

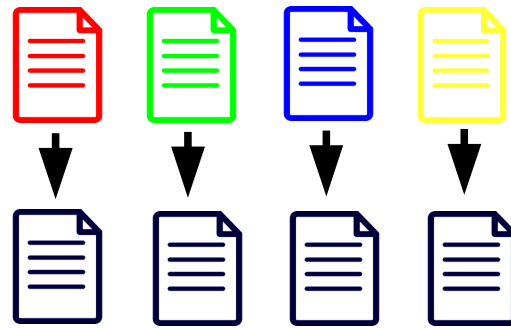
“I'll do the metadata stuff later....”



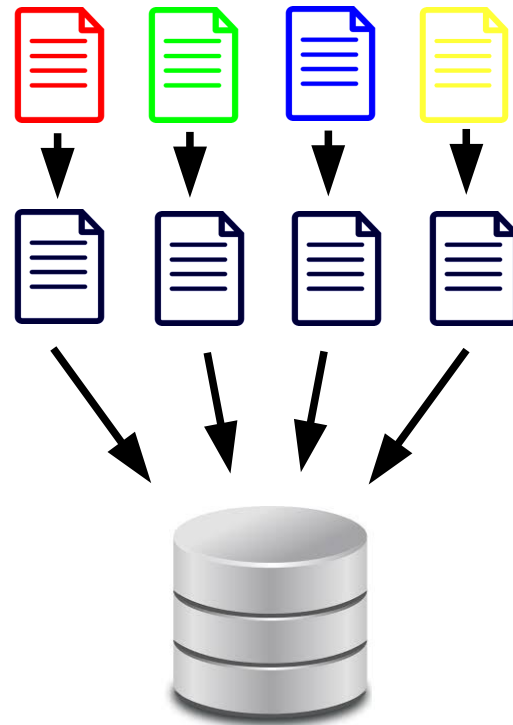
Metadata standards



Consistent format

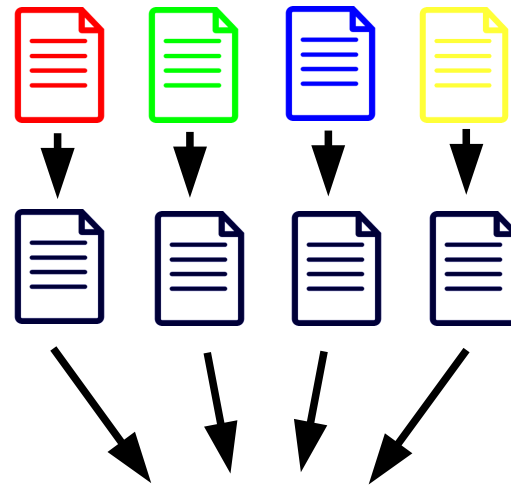


Consistent format



Integration

Consistent format



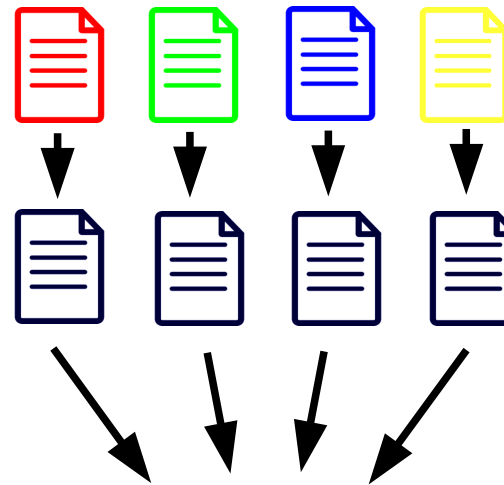
Integration



Searching / Retrieval



Consistent format



Integration

Searching / Retrieval



→ different standards in different research communities



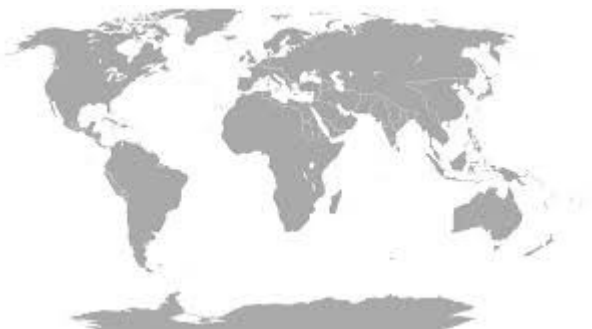
ecology / biodiversity

Ecological Metadata Language (EML)

Darwin Core (DwC)

Access to Biological Collections Data (ABCD)

Biological Data Profile (BDP) of the CSDGM



geospatial

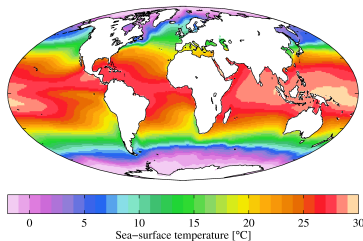
ISO 19115/19139

FGDC Content Standard for Digital Geospatial Metadata (CSDGM)

Directory Interchange Format (DIF)

Open Geospatial Consortium Standards (OGC)

INSPIRE EU directive



oceanographic

geospatial standards and extensions, e.g. Marine Community Profile of ISO 19115

<http://www.oceandatastandards.org/>
(International Oceanographic Data and Information Exchange - IODE)



web resources / media

Dublin Core

Audubon Core

Examples

FixO3: <http://www.fixo3.eu/metadata-catalogue/>

BioFresh: <http://data.freshwaterbiodiversity.eu/metadb/metaDBQry/index.php>

LifeWatchGreece: <http://metacatalogue.portal.lifewatchgreece.eu/>

GBIF: <http://www.gbif.org/dataset>

MedOBIS: <http://ipt.medobis.eu/>

Mercury: <http://mercury.ornl.gov/ocean/>

FixO3: <http://www.fixo3.eu/metadata-catalogue/>

[Simple Search](#) | [Advanced Query](#) | [Help](#)

Search Results: 176 datasets found! (Query time: 0.002 s)

<< [PREV](#) | [1](#) | [2](#) | [3](#) | [4](#) | [5](#) | [6](#) | [7](#) | [8](#) | [9](#) | [10](#) | [NEXT](#) >>

- 1. Maureen Pagnani (2010):** OceanSITES CIS in-situ data (OS_CIS-1_200809_R_O)
Data Center: EUROSITES
Summary: EuroSITES European Ocean Observatory NetworkEU Framework 7 collaborative project contract FP7-ENV-2007-1-202955
Parameters: Temperature; quality flag; sea water electrical conductivity; sea_water_pressure; sea water salinity; mass_concentration_of_oxygen_in_sea_water; sea_temperature_from DOXY sensor; time; Depth of each measurement; Latitude of each location; Longitude of each location
[Data Download](#) - Score: 0.91
- 2. Maureen Pagnani (2010):** OceanSITES CIS in-situ data (OS_CIS-1_200905_O)
Data Center: EUROSITES
Summary: EuroSITES European Ocean Observatory NetworkEU Framework 7 collaborative project contract FP7-ENV-2007-1-202955
Parameters: Temperature; quality flag; sea water electrical conductivity; sea_water_pressure; sea water salinity; mass_concentration_of_oxygen_in_sea_water; sea_temperature_from DOXY sensor; time; Depth of

[Home](#)[About](#)[Summary](#)[Downloads](#)[Versions](#)[Rights](#)[GBIF Registration](#)[Keywords](#)[Contacts](#)[Geographic Coverage](#)[Taxonomic Coverage](#)[Temporal Coverage](#)[Sampling Methods](#)[Additional Metadata](#)

Benthic communities in Amvrakikos Wetlands: Mazoma, Tsopeli, Tsoukalio, Rodia and Logarou lagoons (September 2010 – July 2011)

Latest version published by Hellenic Center for Marine Research on Mar 30, 2016

Zoobenthic communities across confinement gradient in 5 coastal lagoons (Mazoma, Tsopeli, Tsoukalio, Rodia and Logarou) in Amvrakikos for four seasons (data analysis is still in progress)

[GBIF](#)[DwC-A](#)[EML](#)[RTF](#)[Versions](#)[Rights](#)

Downloads

Download the latest version of the resource data as a Darwin Core Archive (DwC-A) or the resource metadata as EML or RTF:

Data as a DwC-A file

[download](#) 303 records in English (79 KB) - Update frequency: unknown

Metadata as an EML file

[download](#) in English (14 KB)

Metadata as an RTF file

[download](#) in English (12 KB)

Versions

Metadatabase view - dataset: **Altmuehl Schwarzach**

▼ General information

dataset entry ID:	UFO_04
name of the dataset:	
full name of the dataset:	<i>Altmuehl Schwarzach</i>
type of dataset (more information):	
data type:	<i>species (taxonomic group) per site database including environmental information</i>
short description of the dataset/summary:	<i>point data/observation data</i>
	<i>two similar restoration projects along the Altmuehl and a tributary (Schwarzach): re-meandering of a slow flowing river, widening of the shoreline, creation of temporary flood channels and backwaters, additional restoration of the floodplain: grassland for ground breeders (Wiesenbrueter), floodplain forest</i>
science keywords according to GCMD:	
topic:	<i>Biosphere</i>
ISO topic category according to ISO 19115:	
	<i>Biota</i>

► Technical and administrative specifications

► Intellectual property rights and citation

► General data specifications

► Site specifications

► Climate and environmental data

► Biological data

► Sample specifications/sample resolution

► Other specifications

GO TO ...

- Login
- Overview of all metadatabase entries
- Metadatabase query tool
- Metadatabase full text search
- PDF-List of all metadatabase entries

EXPORT

- PDF-Export
- EML-Export

HELP

- About the Freshwater Metadatabase
- How to get a login
- How to fill the Metadatabase
- How to query the Freshwater Metadatabase
- About the Freshwater Metadata Journal

Flora-On: occurrence data of the flo...

Occurrence dataset published by Sociedade Portuguesa de Botânica

253,507

Occurrences

[View occurrences](#)

Information

Stats

Activity

Summary

FULL TITLE

Flora-On: occurrence data of the flora of mainland Portugal

DESCRIPTION

This dataset results from the data gathered in the platform Flora-On (<http://www.flora-on.pt/>), which contains observation records of vascular plants within mainland [Portugal](#), made by an array of collaborators. Observations are uploaded directly by collaborators, who are either experienced botanists/naturalists, or occasional collaborators. In this latter case, the observations are first validated by the Editorial Board before publication. The Editorial Board is committed to ensure that only the most rigorous data is published, and it does not release observations whose uncertainty is judged to be high. Flora-On was created by and is kept up exclusively through the voluntary work of associate members of Sociedade Portuguesa de Botânica (Botanical Society of [Portugal](#)), its maintenance costs being supported only by the membership fees of the associates.

TEMPORAL COVERAGES

Date range: Jan 1, 1995 - Feb 9, 2016

LANGUAGE OF METADATA

ENGLISH

LANGUAGE OF DATA

PORTUGUESE

ADMINISTRATIVE CONTACT

METADATA AUTHOR

ORIGINATOR



DOI [doi:10.15468/gi6aum](https://doi.org/10.15468/gi6aum)

PUBLISHED BY

[Sociedade Portuguesa de Botânica](#)

PUBLICATION DATE

Feb 19, 2016

REGISTRATION DATE

May 4, 2015

SERVED BY

[Flora-On datasets](#)

LINKS

• [Dataset homepage](#)

ALTERNATIVE IDENTIFIERS

• [Source webpage](#)

<http://flora-on.pt:8080/intro>

Mercury: <http://mercury.ornl.gov/ocean/>

Dataset_ID: Time_Series_BATS_1994_1996

Investigator(s) Name: Bates, Nicholas

Organization: Bermuda Institute of Ocean Sciences

Address: Bermuda Institute of Ocean Sciences St. George's, GE01, Bermuda

Phone: 441-297-1880

Email: nick.bates@bios.edu

Dataset Info: Dataset ID: Time_Series_BATS_1994_1996

Submission Initial Submission: 20160107

Dates: Revised Submission: 20160107

Cruise: Mooring ID: Time Series

Survey/Experiment Type: Time Series

Survey/Experiment Name: Time_Series_BATS_1994_1996

Cruise ID: BATE19940101

Cruise Info: Time_Series_BATS_1994_1996

Section:

Geographical Coverage:

Geographical Region: Atlantic Ocean

Ports of Call:

Bounds:

North West South East
31.2 -75.4 31.5 -75.6 [Locate](#)

