

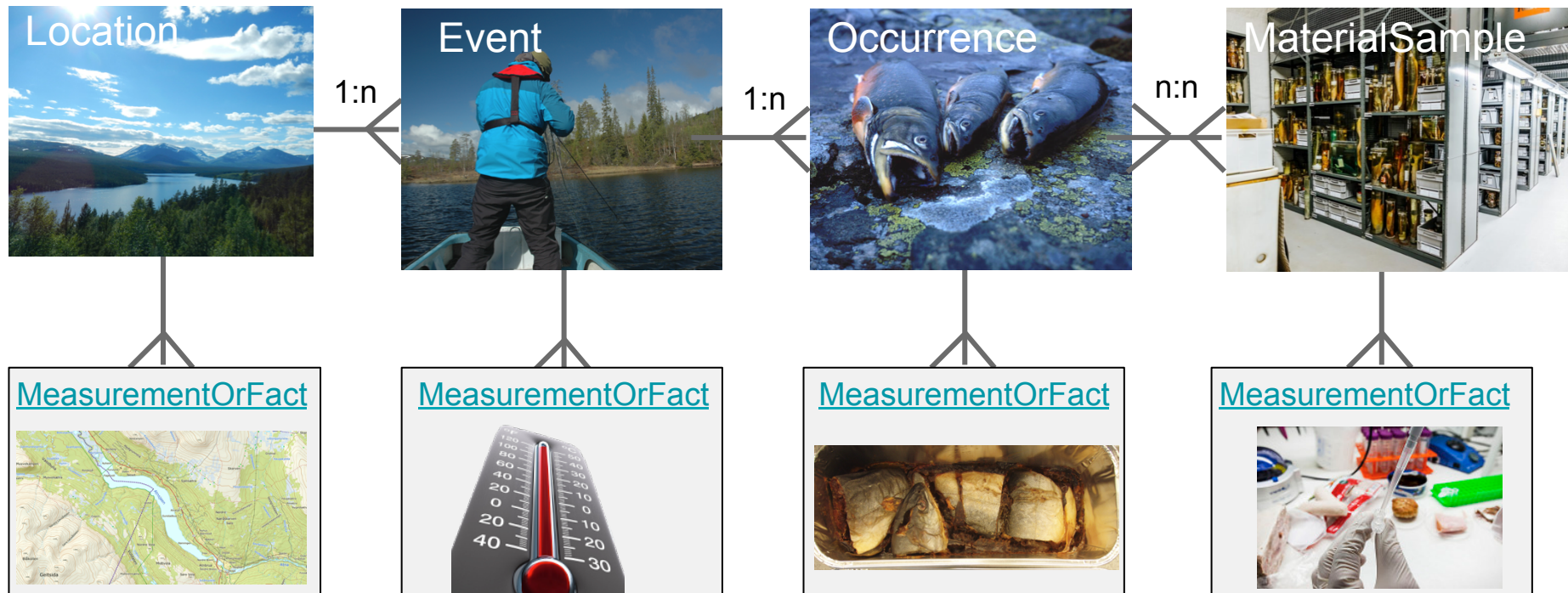
Exploring the use of the event-core and eMoF: Current usage and issues

Anders G. Finstad & Dag Endresen

Content

- The view of the world (biodiversity data in a nutshell)
- Examples of earlier workarounds
- Event-core and eMoF examples and experiences
- Some recommendations

Biodiversity data in a nutshell

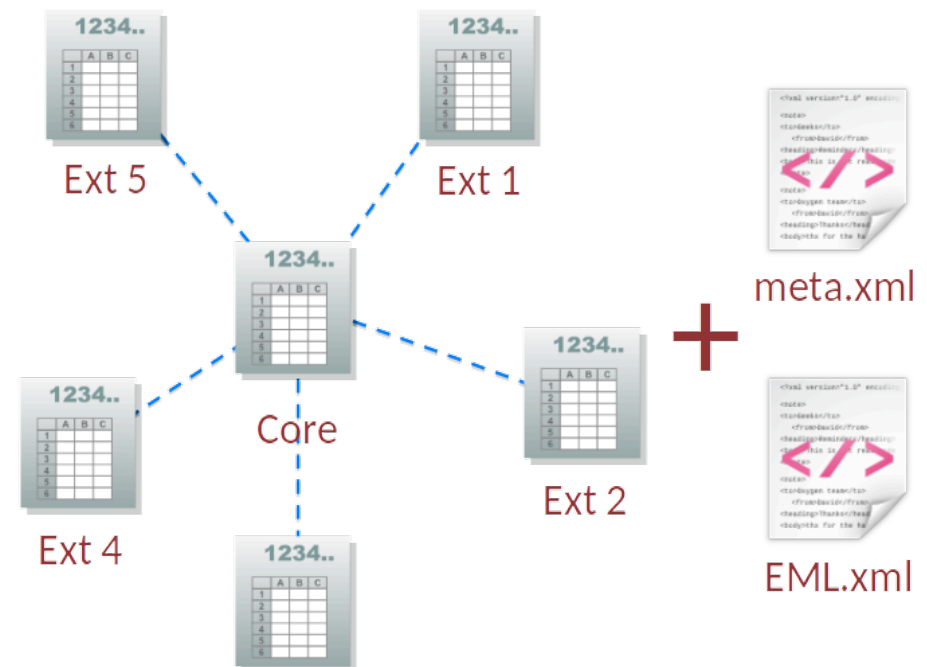


Also, hierarchical structure within different classes of data

Biodiversity data squeezed the data into a DwC-A

Well known issues:

- GBIF only supports “taxon”, “event” and “occurrence” core
- MoFs from other than core not supported



Earlier workarounds to get both
occurrence MoF and events
published



Example: events and occurrences as separate datasets

The image shows two screenshots of a GBIF dataset page. The top screenshot is for the 'Events' dataset, which has 0 occurrences. The bottom screenshot is for the 'Occurrences' dataset, which has 5,701 occurrences. Both datasets are titled 'Gillnet test fishing and radiocaesium (Cs-137) of brown trout (Salmo trutta) and Arctic char (Salvelinus alpinus) 2008 and 2009 from 20 Central Scandinavian lakes:'. The publisher is 'NTNU University Museum' and the contact information is 'Anders Gravbrøt Finstad • Ole Kristian Berg • Eva Ulvan • Ingeborg Helland • Ola Ugedal'. The page includes a green header with a heart icon, a search bar, and a menu icon. The 'Occurrences' dataset has a green button that says '5,701 OCCURRENCES'.

SAMPLING EVENT | REGISTERED 12 JANUARY 2017

Gillnet test fishing and radiocaesium (Cs-137) of brown trout (*Salmo trutta*) and Arctic char (*Salvelinus alpinus*) 2008 and 2009 from 20 Central Scandinavian lakes:

Events

Published by [NTNU University Museum](#)

Anders Gravbrøt Finstad • Ole Kristian Berg • Eva Ulvan • Ingeborg Helland • Ola Ugedal

0 OCCURRENCES

DATASET STATS ACTIVITY DOWNLOAD

OCCURRENCE DATASET | REGISTERED 12 JANUARY 2017

Gillnet test fishing and radiocaesium (Cs-137) of brown trout (*Salmo trutta*) and Arctic char (*Salvelinus alpinus*) 2008 and 2009 from 20 Central Scandinavian lakes:

Occurrences

Published by [NTNU University Museum](#)

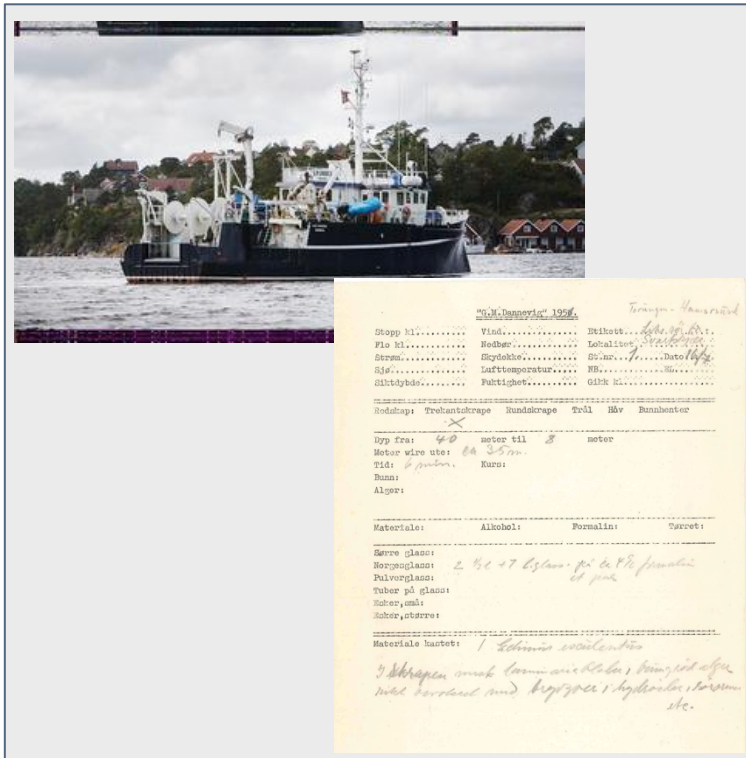
Anders Gravbrøt Finstad • Ole Kristian Berg • Eva Ulvan • Ingeborg Helland • Ola Ugedal

5,701 OCCURRENCES

DATASET STATS ACTIVITY DOWNLOAD

- Requires the use of GUIDs (we'r using urn:uuid)
- Awkward for data user - data not linked by aggregator (GBIF)

Separate datasets needed when material and result from one event are curated independently

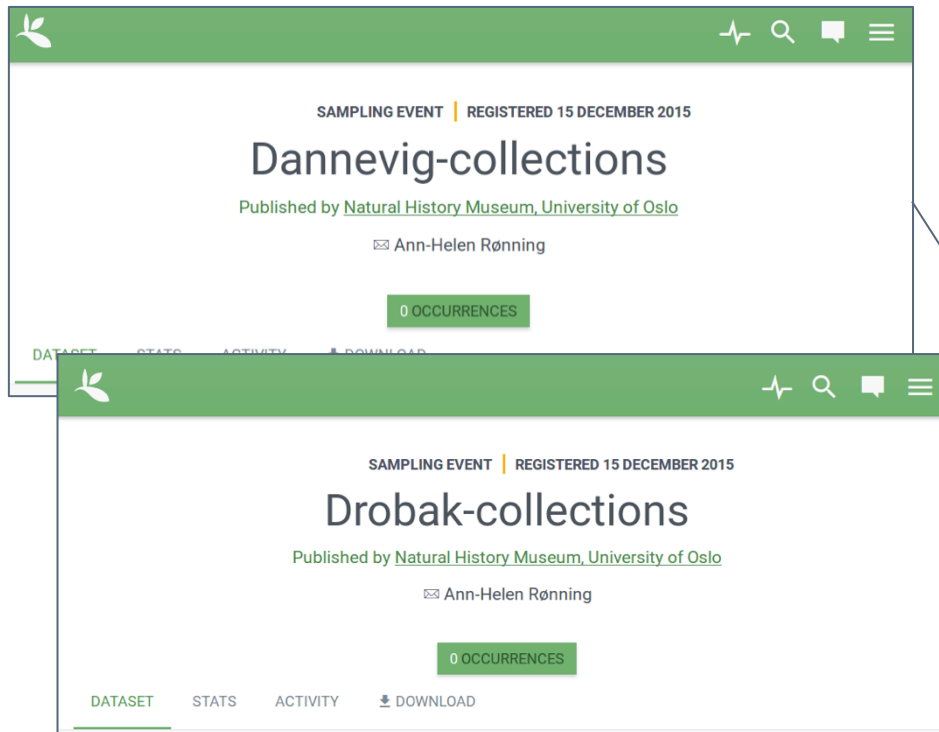


Sampling events

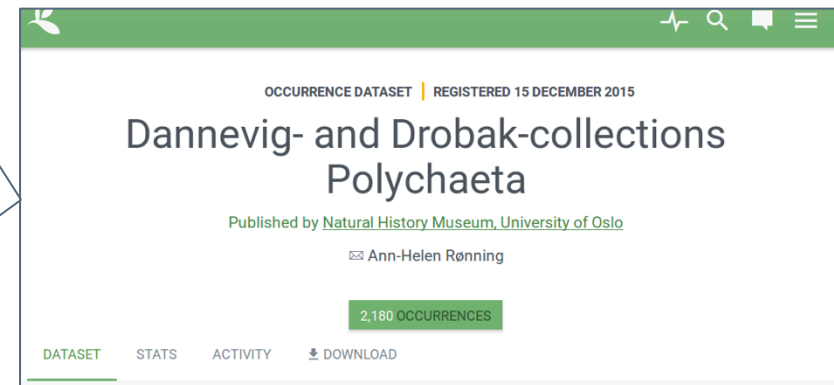


Specimens

Events and occurrences as separate datasets needed when sampling results and material are distributed



Sampling events

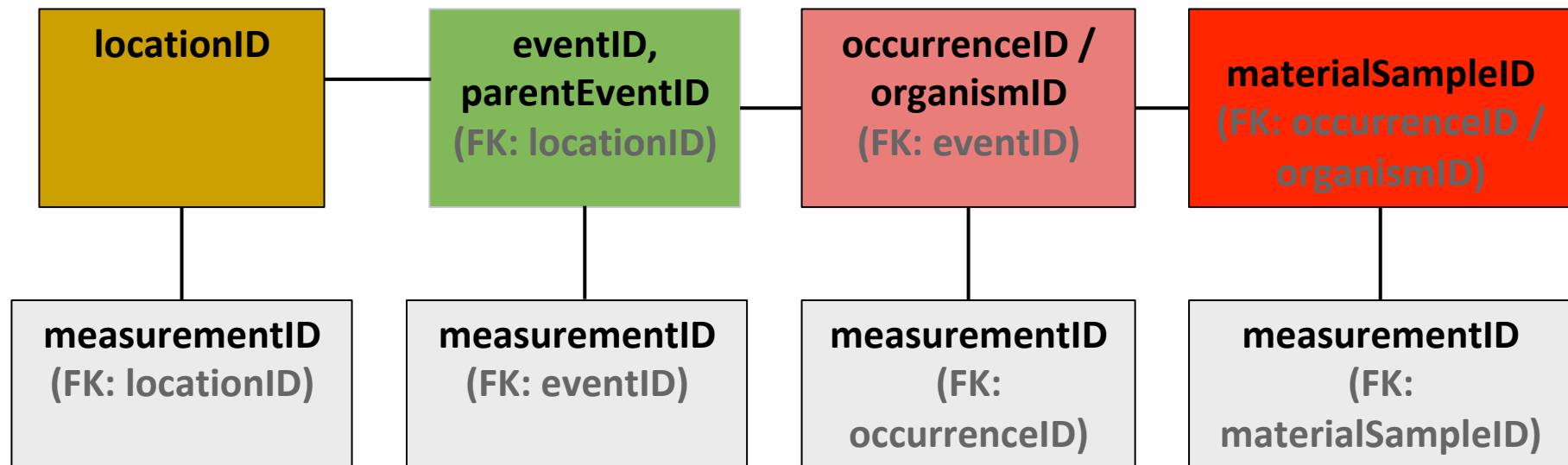


Specimens

Interconnected data published in different datasets is
a generic situation for many datatypes



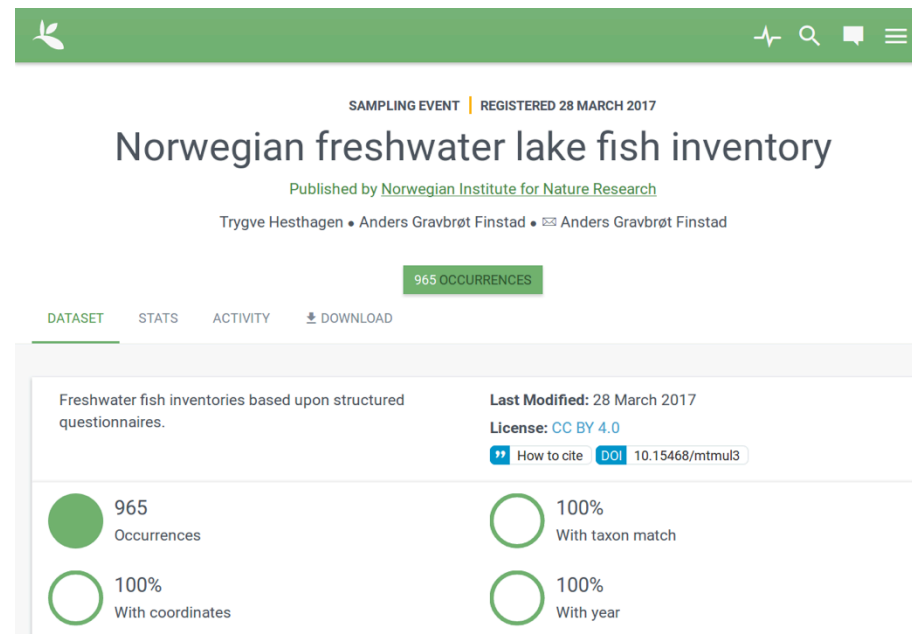
A call for development of the data model and how it's handled by data-aggregators



Information linked together by persistent GUIDs: can be stored and distributed independently

In the meantime; eMoF may provide solutions

- Preliminary user experience good from data-publishing side
- Flexible - handles most traditional ecological data-types
- However, required data wrangling yield elevated user-threshold



Current drawbacks with eMoF

- **Data not returned by GBIF!**
User need to go back to the original DwC-A
- More focus on user driven controlled vocabularies/ ontologies needed (not unique for eMoF)
- Could be made more generic
- i.e. covering other extensions

Record		Occurrence	
Basis Of Record	HUMAN_OBSERVATION	Establishment Means	UNCERTAIN
Dataset Name	Freshwater lake fish inventory	Individual Count	
		Life Stage	
		Occurrence ID	urn:uuid:8a79031e-baa6-433b-8d39-d634ac18b941
		Occurrence Status	irregular
		Organism Quantity	NA
		Organism Quantity Type	NA
		Record Number	skjemaNr: TH_NL_2015_88 art: 5 Inr: 273 innført_år:
		Recorded By	Trygve Hesthagen
		Reproductive Condition	NA
		Sex	
Event		Taxon	
Year	2015	Kingdom	Animalia
Event Date	2015-01-01T00:00:00.000+00:00	Phylum	Chordata
Event ID	urn:uuid:42bd004b-ef79-4962-94a6-17ea58c92413	Class	Actinopterygii
Field Number	skjemaNr: TH_NL_2015_88 innført_år:	Order	Salmoniformes
Parent Event ID	NA	Family	Salmonidae
Sample Size Unit	NA	Genus	Salmo
Sample Size Value	NA	Specific Epithet	trutta
Sampling Effort	NA	Scientific Name	Salmo trutta Linnaeus, 1758
Sampling Protocol	survey_questionary	Scientific Name Authorship	
		Rank	SPECIES
		Vernacular Name	Brown trout
Location		Other	
Coordinate Uncertainty In Meters	15	Identifier	urn:uuid:42bd004b-ef79-4962-94a6-17ea58c92413
Country	Norway	Record License	http://creativecommons.org/licenses/by/4.0/legalcode
Country Code	NO	Modified	2017-03-25T00:00:00.000+00:00
County	Oppland		
Decimal Latitude	61.193574		
Decimal Longitude	9.021414		
Geodetic Datum	WGS84		
Georeference Remarks	Latitude/longitude coordinates represent waterBody centroid		
Location ID	NVE Vatn_Innr:32999 NOFA locationID:e9376720-849f-49b8-8b43-53d7d2a957a9		
Municipality	Øystre Slidre		
Water Body	nordre Vindin		
CITATION			
Hesthagen T, Finstad A G (2017). Norwegian freshwater lake fish inventory. Norwegian Institute for Nature Research. Sampling_event Dataset https://doi.org/10.15468/mtmul3 accessed via GBIF.org on 2018-01-14. https://www.gbif.org/occurrence/1450645797			

Recommendations

Short term (eMoF and DwC-A)

- Urgent need for better handling of event data and MoF/eMoF data in GBIF dataproducts
- Low user-threshold data publisher tools (data-wrangling)
- Stimulate usage of controlled vocabularies / ontologies; e.g. revitalize terms.tdwg.org
- Documentation and good usage examples lacking

Long term

- Need for a more flexible data-model ([see previous slide](#)) replacing or building on the DwC-A
- Data aggregators such as GBIF must pull data together from several datasets (use of GUIDs should be obligatory for core terms)
- Usage of controlled vocabularies / ontologies (as for short term)