Management of Marine Media Metadata

GEOMAR data management team: Carsten Schirnick & Claas Faber Pina Springer, Hela Mehrtens, Lisa Paglialonga





Management of Marine Media Metadata

from a data management perspective



Metadata = crucial information for research

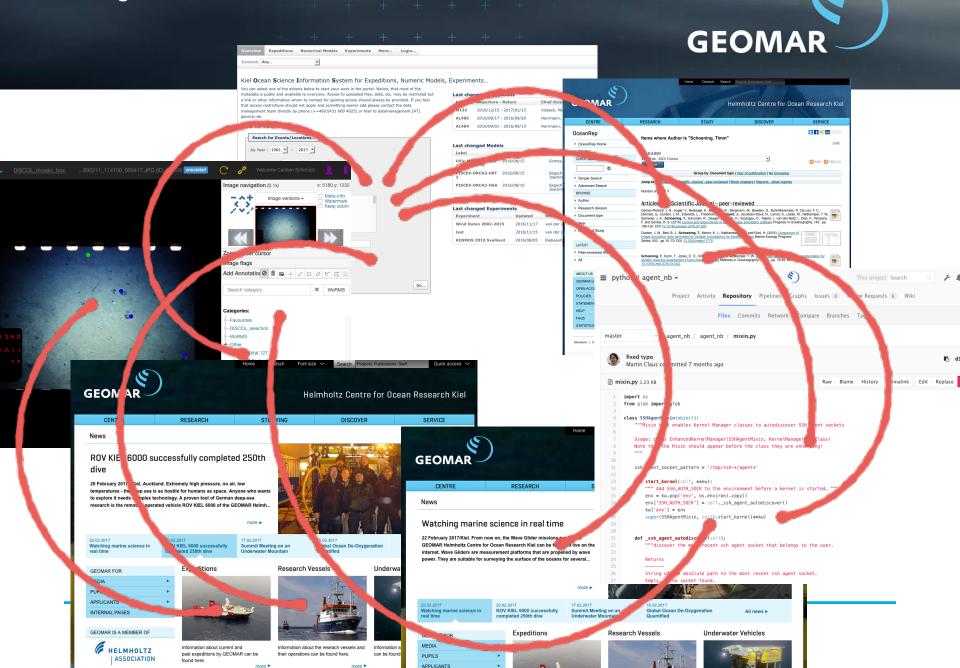
- location (time, position, depth)
- context (topic, project, cruise, platform, ...)
- related information (temperature, salinity, oxygen,...)
- provenance (processing steps, version, status, ...)
- and annotations...?



... metadata add value to research data

- for processing, e.g. by recording
 - file versions
 - authors, contributors
 - processing steps and applications
 - quality control
- for management of research data
 - enable searching for content
 - link with projects, expeditions, publications,...

Management of Marine Media Metadata





... media metadata are heterogeneous

- across institutions, projects, topics and disciplines
- but also within institutions
 - between instruments and recording devices
 - between operating teams
 - between responsible scientists and data managers
- and even by individual researchers...



... metadata in the wild - hard to handle

- File: 20150323_S0239_lower_HD_029ROV02.mov
- Folder structure:
 - /var/data/media/ROV-KIEL6000/ 2015-03-10_SO239/SO239_029ROV02/upper_HD/...
- EXIF comment field
- image overlays
- separate containers (.txt, .csv, .zip, .xlsx, ...)
- in management tools and media servers (=databases)



Problems

- heterogeneity
 - prevents easy discovery and
 - hinders exchange
- metadata storage is often not robust
 - overwritten file names, typos, moving to new storage
- metadata not machine readable and needs interpretation
 - prevents discovering and searching across systems
- ... where are the metadata ?



- add metadata to the media assets
- unified metadata standard for marine media (Exif + XMP)
 - agreement on common fields, formats, structure*
 - interoperable technical solutions
 - make media file primary source of metadata

(* agree on common base set of fields and provide process for specific extensions)



- XMP as extension of Exif
 - ISO standard for attaching metadata to files
 - mature, implemented for many file types
 - Extensible using namespaces

- but how to record current version of metadata
- preserve history of metadata
- how to disseminate latest metadata



- proper location for preservation of metadata records
 - web based services
 - longterm accessibility
 - must be interoperable

- persistent identifier assigned to asset (e.g. handle)
 - unique identification
 - makes citable and permanently findable



- best practice for persistent identifiers (e.g handle, DOI)
 - lead to landing page
- landing page aggregates
 - metadata and their history
 - access methods to asset and related data
 - re-usage and citations
- auto generated but human- and machine readable
 - trusted single-point-of-truth always up-to-date



- discuss in community
- find support & cooperation
- Research Data Alliance (RDA) ?



...thanks