Digital Cultural Heritage Thematic Services of the Open Call

NI4OS-Europe Open Call Training Event
14 April 2022

Dr. Georgios Artopoulos
Assist. Professor
The Cyprus Institute
&
Maria Tzima & Iason Giraud
Research Assistant
The Cyprus Institute

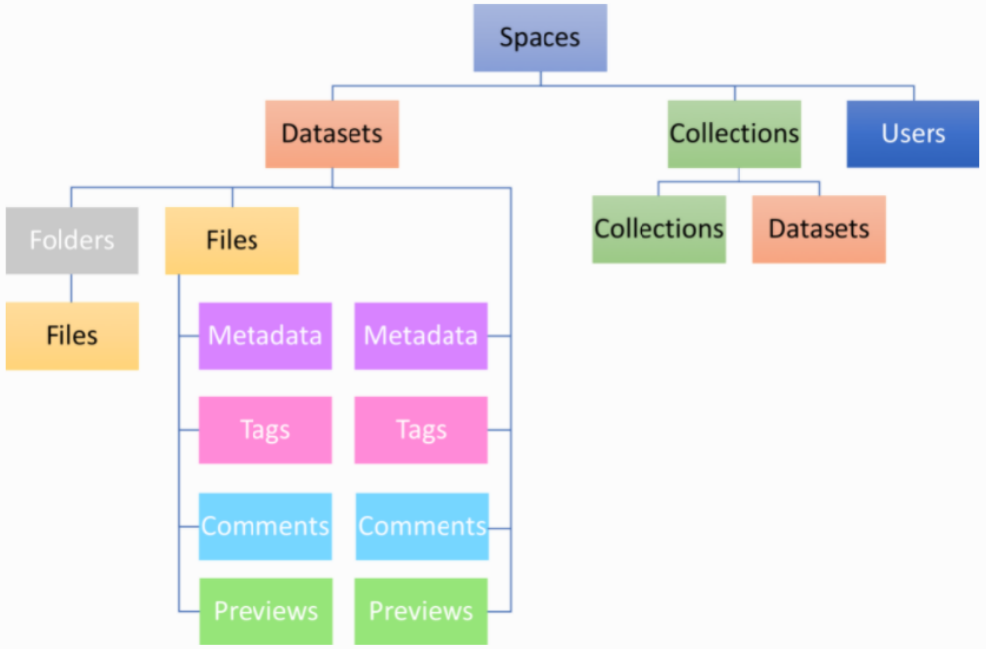
Introduction to Clowder4DCH

Platform as a service aimed at researchers and educators

What is Clowder4DCH?

- The content management system for the Cultural Heritage communities of the Horizon 2020 funded NI4OS project (https://ni4os.eu/).
- Is a highly extensible active curation-based research data management platform.
- Target Users: GLAM industry (galleries, libraries, archives and museums) and Education
- Access Mode: Open, with user password authentication.

Data Model



C4DCH helps manage data by organizing files and metadata in *folders*, *datasets*, *collection*s and *spaces*.

Users

Users can have different roles

- Viewers: can view and download data
- Editors: viewers + manage (add, remove) data in datasets, collections and spaces
- Administrators: editors + can manage users and roles

FAIR Principles

C4DCH and FAIR Principles

- Findable: Data are described with rich metadata and can be found through search.
- Accessible: Clowder is an open platform. Data and metadata are accessible and retrievable.
- Interoperable: Data and metadata can include references to other data.
- Reusable: Data and metadata can be richly described by comments and attributes. Available data usage license.

Metadata Handling (1/3)

Flexible Metadata Representation

Two types of metadata:

- Required metadata, such as time of creation, author, license, etc.
- Optional generic metadata. The optional metadata can be produced by users and extractors.

Metadata Handling (2/4)

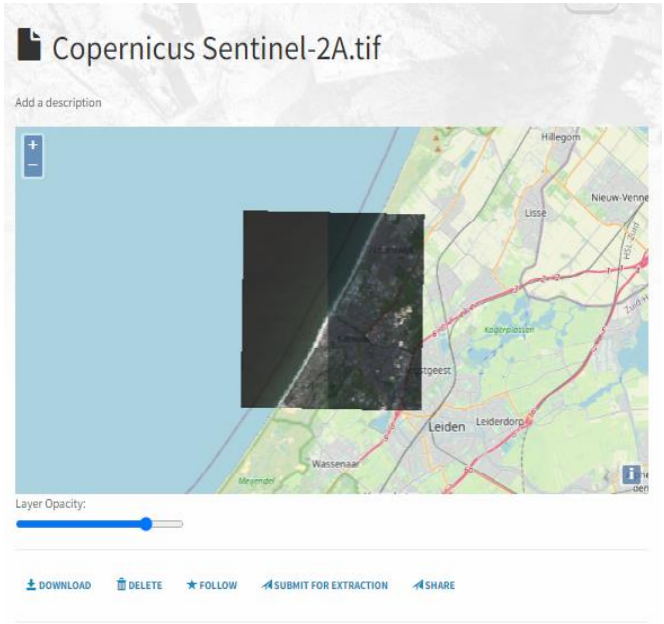
Handling Metadata by C4DCH

- JSON Linked Data (JSON-LD) is used to encode optional metadata.
- The metadata definitions is managed at the Clowder instance level by an administrator and at a space level by space administrators.
- Default metadata definitions include a subset of Dublin Core metadata definitions. Each metadata definition is defined by a label, description, URI identifying the term, and type.

Metadata Handling (2/3)

Automatic Metadata Extraction

When new data is added to the system, whether it is via the web front-end or through its Web service API, a cluster of extraction services process the data to extract interesting metadata and create web based data visualizations.



Metadata Handling (3/3)

What is an Extractor?

Extractors

Extractors are independent processes running outside of the main C4DCH application. Once you have uploaded your files, you can use the extractors in order to visualize your data or extract interesting metadata.

- Separate processing modules that interact with data within C4DCH
- Can be run anywhere via the same message bus - RabbitMQ - as C4DCH
- Extractors download files and upload the results back to C4DCH

List of Available Extractors

- audio preview
- Shapefiles preview
- GeoTiff preview
- image preview
- pdf preview
- video preview
- nlp simple language
- file digest
- cv river
- GeoTiff metadata
- image metadata
- image ocr
- nlp simple summary
- nlp tika

Get in Touch

We'd love to hear your thoughts, feedback and questions

ADDRESS

20 Konstantinou Kavafi Street

2121, Aglantzia, Nicosia, Cyprus

CONTACT

g.artopoulos@cyi.ac.cy

SECURITY CONTACT EMAIL

m.tzima@cyi.ac.cy