



Clowder4DCH

User Guide

Service Provider	The Cyprus Institute
-------------------------	----------------------

Component Provider	The Cyprus Institute
---------------------------	----------------------

Status	Final
---------------	-------

Version and date	[Version and date]
-------------------------	--------------------

	[Service name]
--	----------------

The work was supported by the European Commission through the H2020 Project NI4OS-Europe Contract 857645 [and other acknowledgements] and is licensed under a

[Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)



Contents

1	Introduction	3
1.1	Information	3
2	Accessing the Service	4
2.1	Requirements	4
2.2	Signup Process	4
2.3	Login Procedure	5
2.4	Editing your Profile	6
3	Basic Concepts	7
3.1	Data Model	7
3.2	Users	8
4	Documentation of Modules and Workflows	9
4.1	Datasets	9
4.2	Spaces	13
4.3	Collections	16
4.4	Extractors	19
4.5	Previewers	24
5	Troubleshooting	26



1 Introduction

Clowder4Digital Cultural Heritage is the content management system for the Cultural Heritage communities of the Horizon 2020 funded NI4OS project (<https://ni4os.eu/>). C4DCH is a highly extensible active curation-based research data management platform. Users can link and organize datasets in online collections for the creation of semantically structured data repositories specialized for digital cultural heritage. C4DCH enables users to form an online collaboration environment to support research communities and activities, and disseminate results.

1.1 Information

- Platform as a service.
- Target Users: Museums, Galleries, Archives and Collections.
- Access Mode: Open, with user password authentication.
- Main Characteristics:
 - ⇒ Variety of Supported Data Types
 - ⇒ Flexible Metadata Representation
 - ⇒ Automatic Metadata Extraction
 - ⇒ Data Visualizations



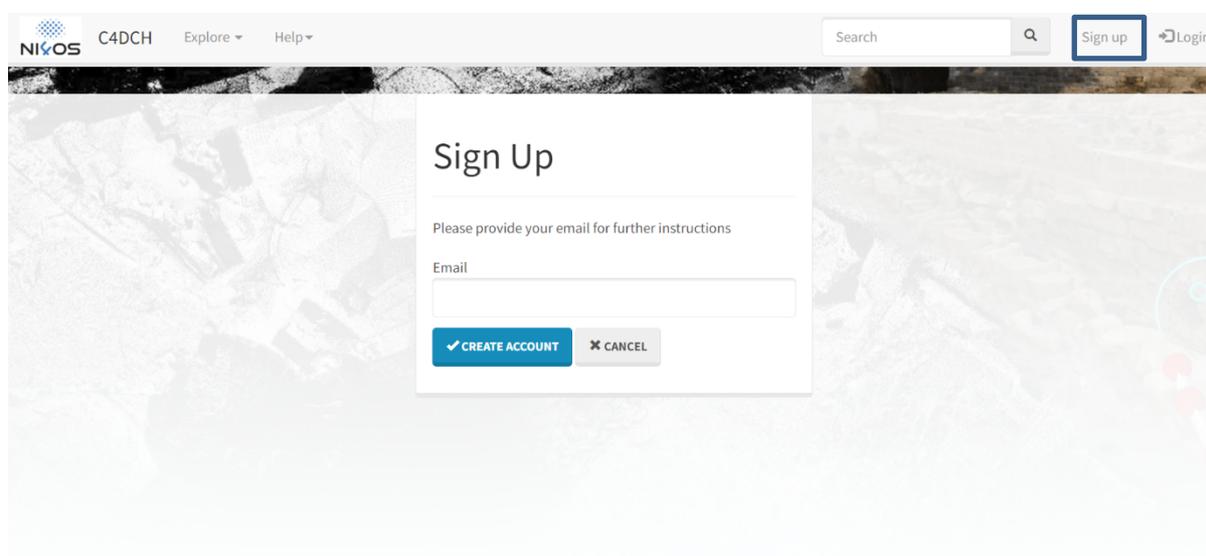
2 Accessing the Service

2.1 Requirements

To access web user interface of this service the user has to use a modern and standards compliant web browser with enabled support for JavaScript, such as the current versions of Google Chrome, Mozilla Firefox, Apple Safari or Microsoft Edge.

2.2 Signup Process

Click on the highlighted link in blue (<https://clowder.hpcf.cyi.ac.cy/>). First step is to complete the registration to the C4DCH platform. Click on the Sign Up button, in the top right of the page. Input your email and create an account. After you input your email and click on Create Account Button, you will receive an email with instructions to continue the signup process. Your account will be activated immediately, and you will receive a confirmation email.



The screenshot displays the C4DCH web interface. At the top, there is a navigation bar with the NIKOS logo, 'C4DCH', 'Explore', and 'Help' menus. A search bar and 'Sign up' and 'Login' buttons are also visible. The main content area features a 'Sign Up' modal form. The form has a title 'Sign Up' and a prompt: 'Please provide your email for further instructions'. Below this is an 'Email' input field. At the bottom of the form are two buttons: 'CREATE ACCOUNT' (with a checkmark icon) and 'CANCEL' (with an 'X' icon). The background of the page is a blurred aerial photograph of a city.

Figure 1 Sign up process



2.3 Login Procedure

Once you are already registered, you can use your credentials to log in to C4DCH. First step is to click on Login Button in the top right section of the screen. You can login with your email and password used in the step above.

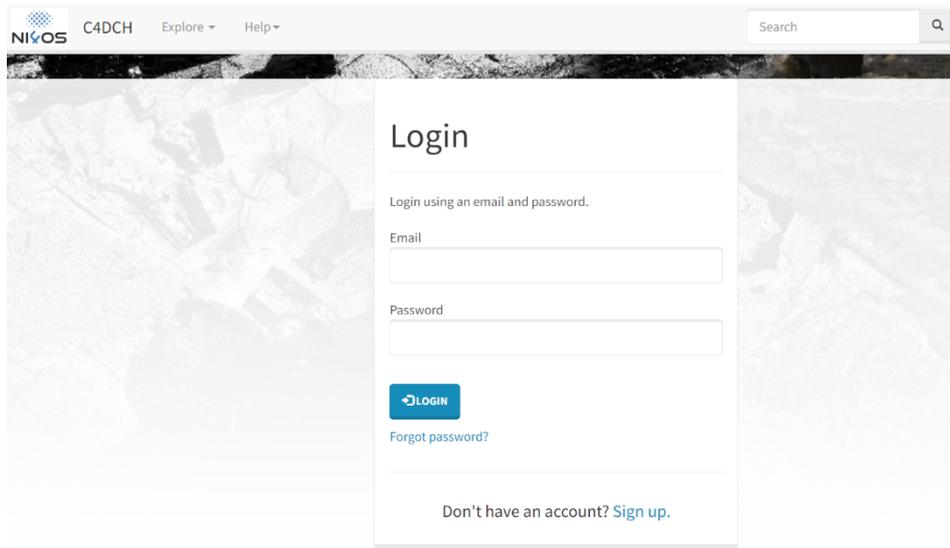
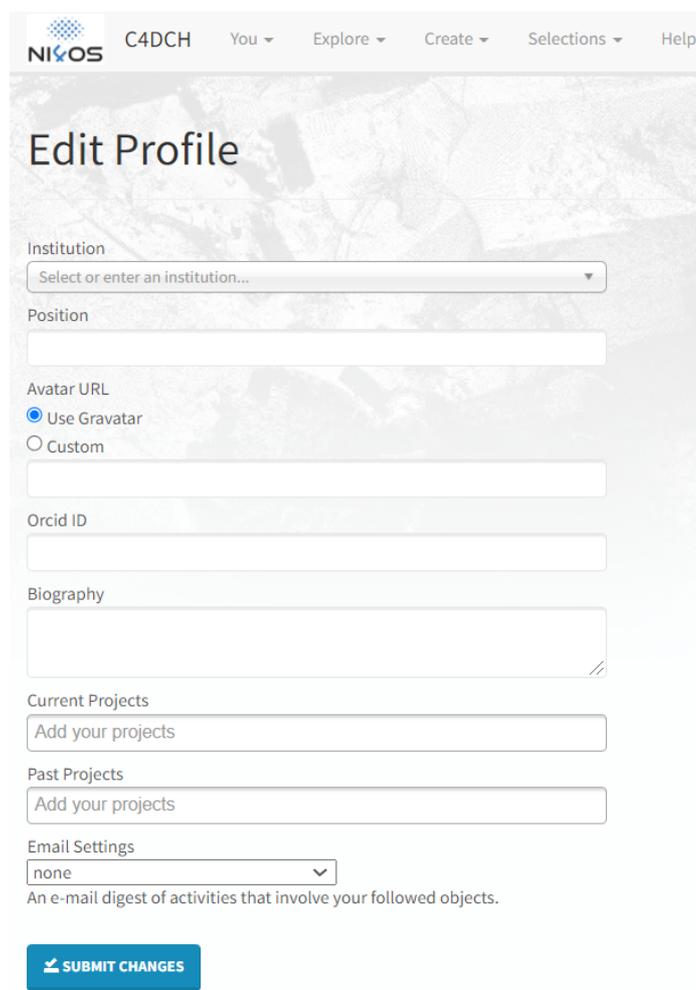


Figure 2 Login Environment



2.4 Editing your Profile

Now you can edit your profile! C4DCH gives you the capability to add personal information to your profile, such as Institution, Position, Biography, Projects etc.



The screenshot shows the 'Edit Profile' page in the C4DCH application. The page has a header with the NIXOS logo and navigation links: C4DCH, You, Explore, Create, Selections, and Help. The main content area is titled 'Edit Profile' and contains several form fields:

- Institution:** A dropdown menu with the placeholder text 'Select or enter an institution...'
- Position:** A text input field.
- Avatar URL:** Radio buttons for 'Use Gravatar' (selected) and 'Custom', followed by a text input field.
- Orcid ID:** A text input field.
- Biography:** A large text area with a 'done' icon in the bottom right corner.
- Current Projects:** A text input field with the placeholder text 'Add your projects'.
- Past Projects:** A text input field with the placeholder text 'Add your projects'.
- Email Settings:** A dropdown menu currently set to 'none', with a note below it: 'An e-mail digest of activities that involve your followed objects.'

At the bottom of the form is a blue button labeled 'SUBMIT CHANGES' with a checkmark icon.

Figure 3 Profile Editing



3 Basic Concepts

Now you are ready to upload and share some data! After logging in C4DCH redirects you in the home page of the platform. The home page is where the site redirects you after logging in. You can navigate back to your home page by clicking the NI4OS logo, or the C4DCH button in the top navigation.

3.1 Data Model

Clowder helps manage data of any type by organizing files and metadata in folders, datasets, collections and spaces.

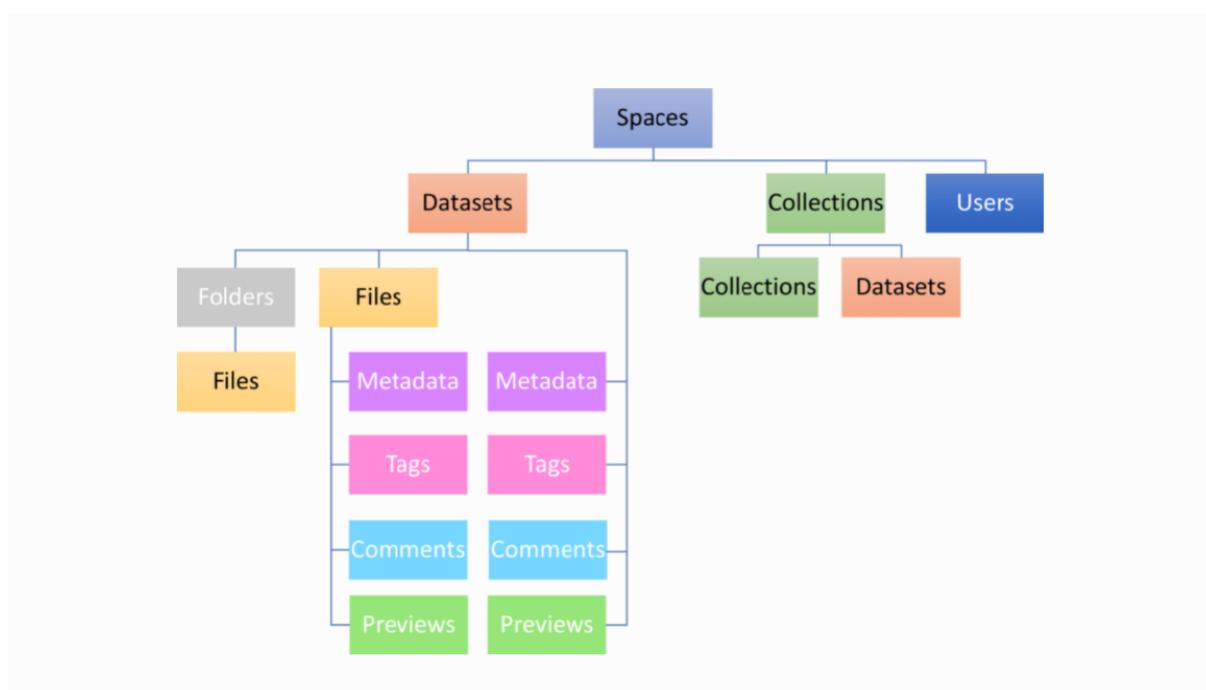


Figure 4 C4DCH Data Model

Datasets are the main resource type. A group of files that through some defined relationship or corresponding metadata are strongly tied together and not representable otherwise by the individual files. Each dataset has a name, description, owner and license.

Collections are a user defined group of datasets and other collections. Collections of datasets can only contain datasets and nested collection but have no generic metadata associated with them.

Spaces are groups of collections, data sets, and files with defined user access rights. Spaces are used to share data within datasets and collections with other users.



3.2 Users

In C4DCH users can have different roles; viewers, editors and administrators.

A viewer can only see the files, datasets, collections and spaces.

An editor, besides the view privileges can manage datasets, collections and spaces space. For example, an editor can add datasets or remove files.

An admin can do what an editor does and also manage users and roles within a space.



4 Documentation of Modules and Workflows

This chapter includes the explanation of the basic functionalities and capabilities of Datasets, Collections and Spaces in C4DCH. It also contains a description of some interesting and helpful C4DCH tools, such as extractors and previewers.

4.1 Datasets

Creating Datasets

Datasets can be created with five different ways. The first option is to create the Datasets from the home page, by clicking the Create Dataset button which is shown in the Figure 5.

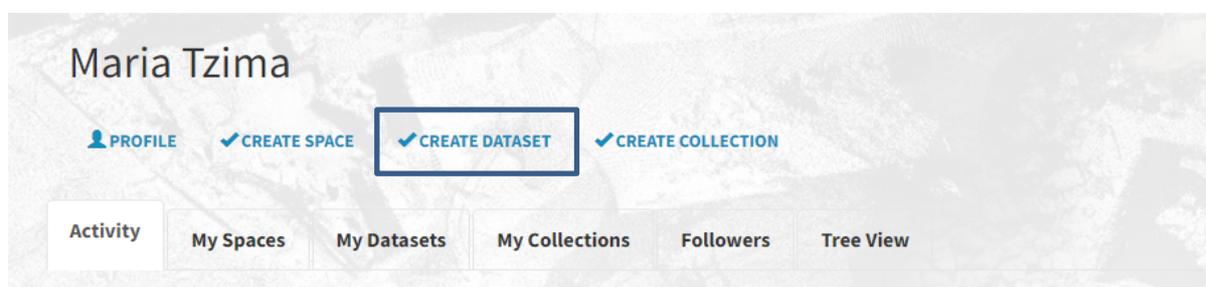


Figure 5 Dataset Creation 1

The second way is to create the Dataset from the top navigation. You have to use the Create button and then select the option Datasets, as it is shown in the Figure 6.

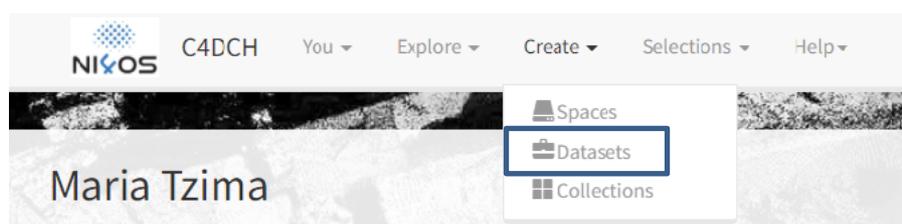


Figure 6 Dataset Creation 2



The third option is to create a new Dataset from the Dataset page. You can enter the Dataset page, by clicking the Explore button in the top navigation. Next step is to select the Create option (Figure 7).

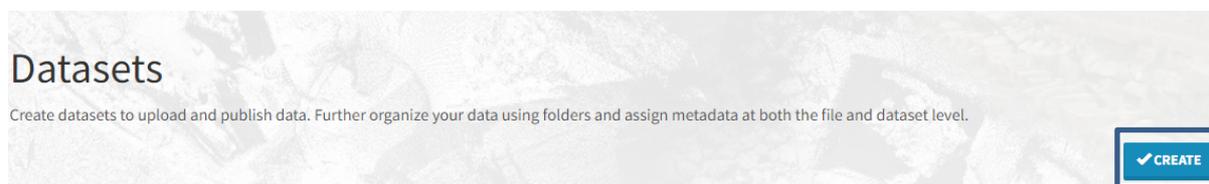


Figure 7 Dataset Creation 3

The next way is to create a Dataset within a Space. In order to do that, you have to enter the Space page from the top navigation and then to click the Create Dataset button (Figure 8).



Figure 8 Dataset Creation 4

The last way is to create Datasets within a Collection. Firstly, you enter the Collection page and then select the Create Dataset button, as can be seen in Figure 9.

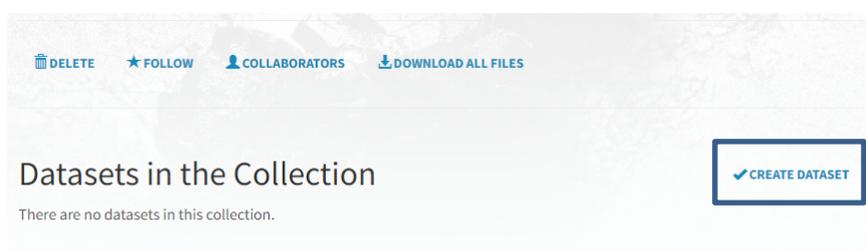
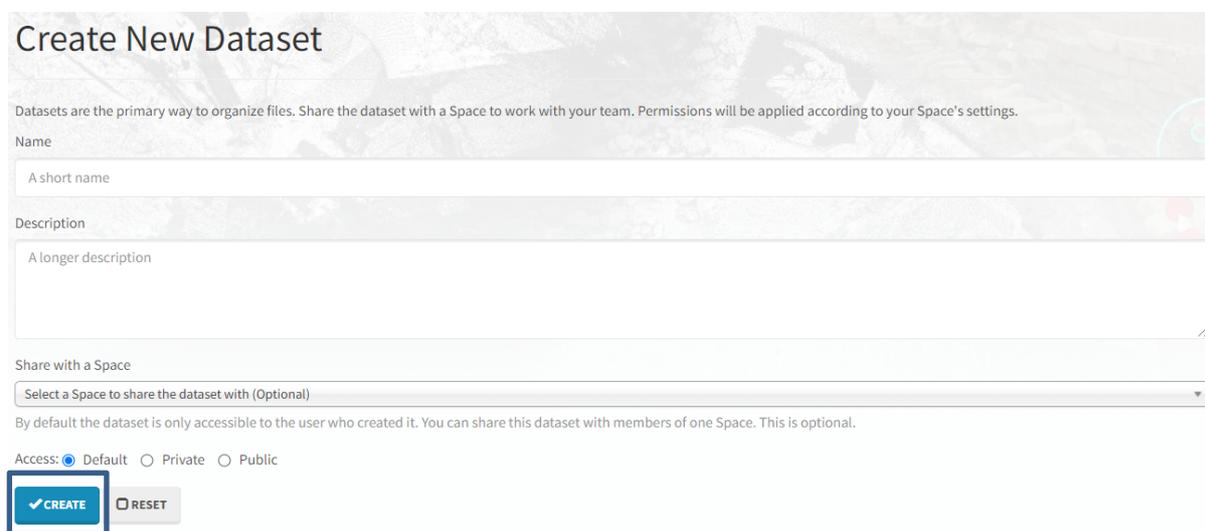


Figure 9 Dataset Creation 5



In order to create the Dataset, you need also to add some basic information about the Dataset, such as Name, Description. The only required field is the Name of the Dataset. When the form is completed, click the Create button (Figure 10).



Create New Dataset

Datasets are the primary way to organize files. Share the dataset with a Space to work with your team. Permissions will be applied according to your Space's settings.

Name
A short name

Description
A longer description

Share with a Space
Select a Space to share the dataset with (Optional)

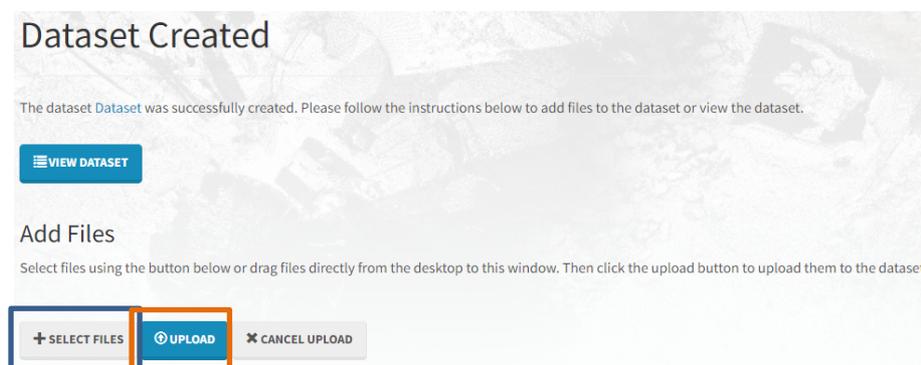
By default the dataset is only accessible to the user who created it. You can share this dataset with members of one Space. This is optional.

Access: Default Private Public

Figure 10 Dataset Creation 6

Uploading Files

After creating the dataset, you can add files with two different ways (Figure 11). The first method is to Drag files to the interface and then select the Upload button. The second way is to click on the Select Files button, which opens a navigation page within your system to select the files you want to upload, and then click the upload button.



Dataset Created

The dataset Dataset was successfully created. Please follow the instructions below to add files to the dataset or view the dataset.

Add Files

Select files using the button below or drag files directly from the desktop to this window. Then click the upload button to upload them to the dataset.

Figure 11 Uploading Files



Additional Capabilities

Within a dataset are available a lot of capabilities such as:

- Creating Folders

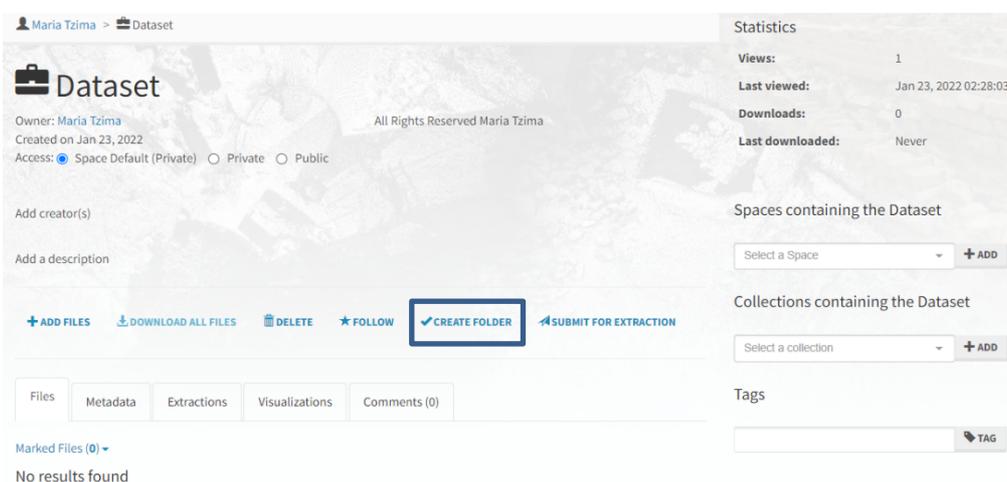


Figure 12 Creating Folders

- Moving Files
- Adding to a Space
- Removing a Dataset from a Space
- Adding Metadata

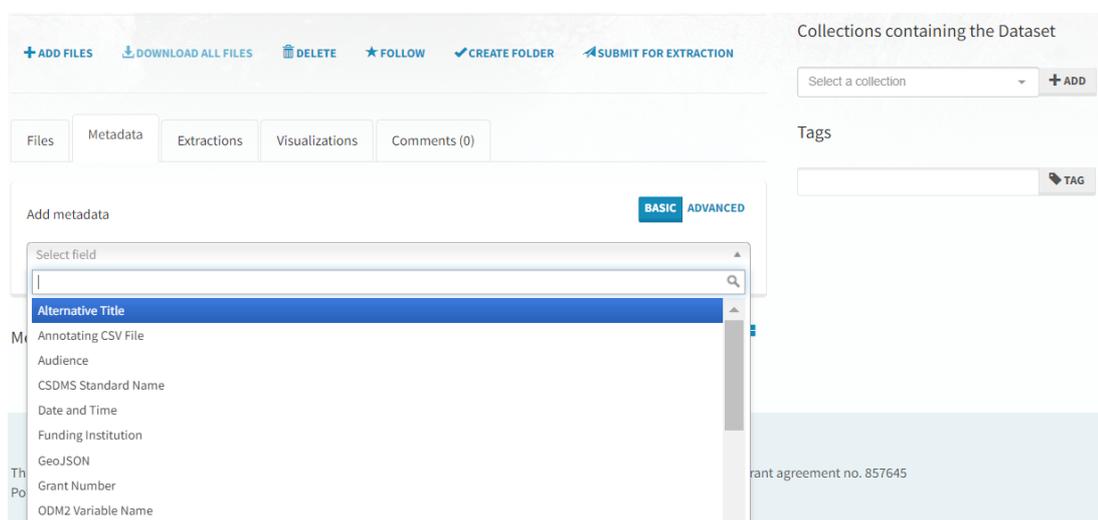


Figure 13 Adding Metadata

- Adding Tags



4.2 Spaces

Creating Spaces

Spaces can be created with three different ways. The first way is to create a Space from the top navigator. Click on the Create dropdown in the top navigation and then choose Spaces (Figure 14).

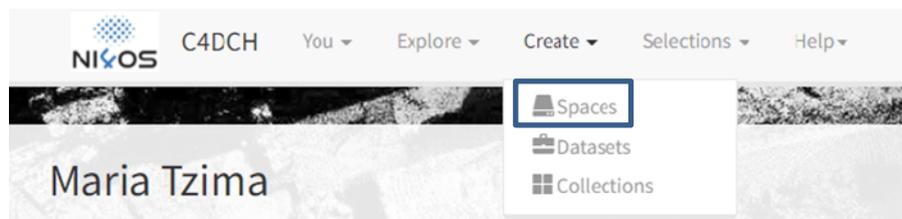


Figure 14 Creating Spaces 1

The second way is to create the Space from the home page. Click on the Create Space button (Figure 15).

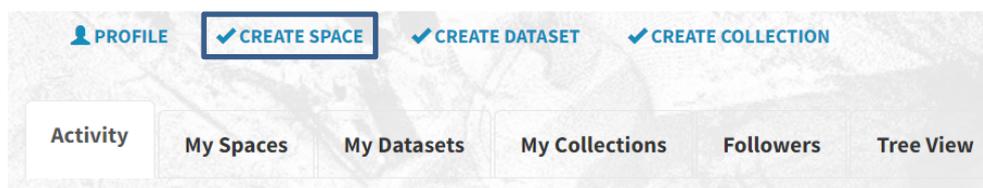


Figure 15 Creating Spaces 2

The third method is to create the Space from the Space list. You can enter the Space page, by clicking the Explore button in the top navigation. Next step is to select the Create option (Figure 16).



Figure 16 Creating Spaces 3



In order to Create a new Space, the only required field is the Name of the Space. You can also add a description, external links, change the access etc. When the form is completed, click the Create button (Figure 17).

Create New Space
Spaces are collaborative workspaces that contain people (with designated permissions), branding, and Datasets.

Name

Description

External Links
 [add](#)

Logo URL

Recommended image size is 100 x 100

Banner URL

Recommended image size is 1170 x 200

Affiliated Spaces
 [add](#)
The names or IDs of other spaces whose published data should be shown on this space's Published Data page

Access
 PRIVATE PUBLIC
This sets the privacy level for collections, datasets, and files in the space. Access can be overwritten at the individual dataset level. By marking access "Public" you make the space's content available to users outside of your project team.

Figure 17 Creating Spaces 4



Additional Capabilities

Within a dataset are available a lot of capabilities such as:

- Editing a Space
- Inviting Users and Adding Users to a Space

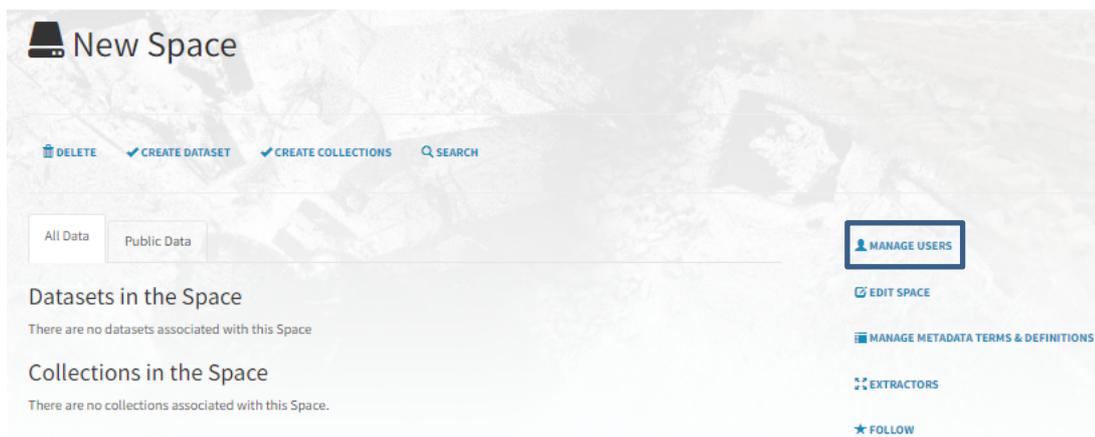


Figure 18 Inviting Users

- Requesting and Granting Access to a Space

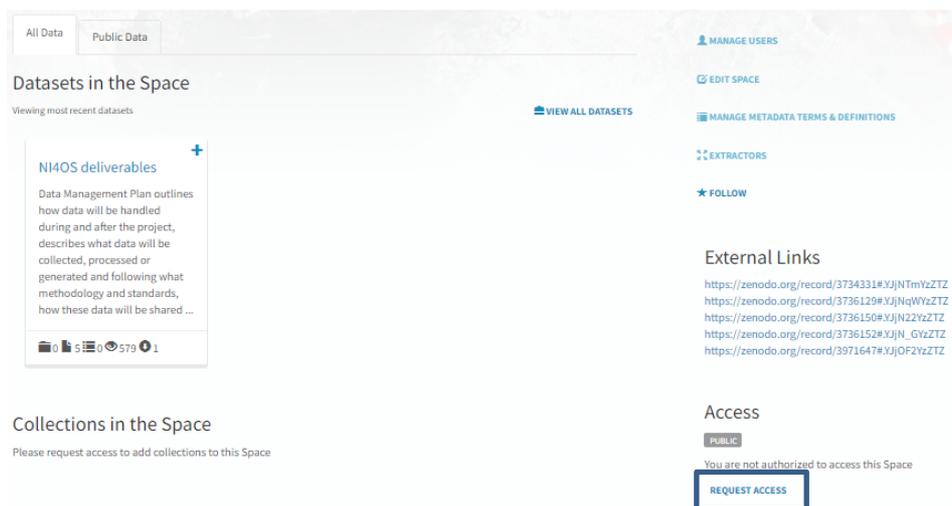


Figure 19 Requesting Access

- Deleting a Space



4.3 Collections

Creating Collections

Collections can be created with four different ways. To create a collection from the home page, click on the Create Collection button, displayed in blue in Figure 20.



Figure 20 Creating Collections 1

The second way is to create a collection from the top navigator. Click on Create button and then select the option Collections. It can be shown in Figure 21.

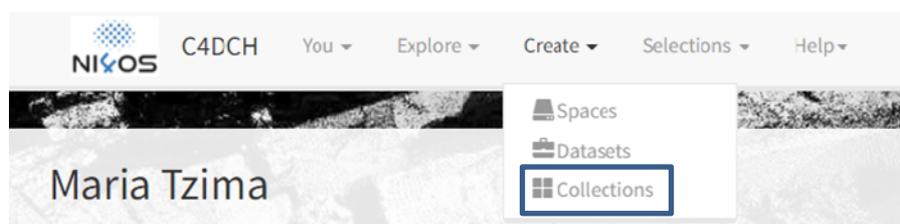


Figure 21 Creating Collections 2

The third method is to create the Space from the Space list. You can enter the Collection page, by clicking the Explore button in the top navigation. Next step is to select the Create option (Figure 22).

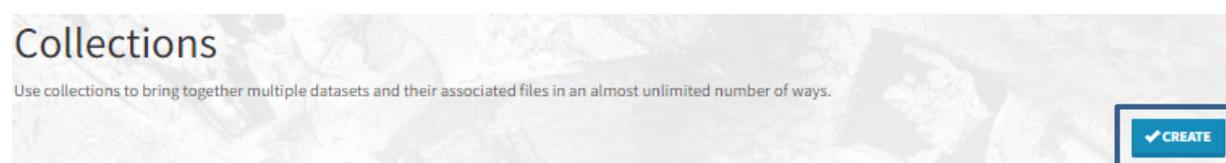


Figure 22 Creating Collections 3



The last method is to Create a Collection within a Space. Firstly, you need to enter the Space and then click on the Create Collections button (Figure 23).

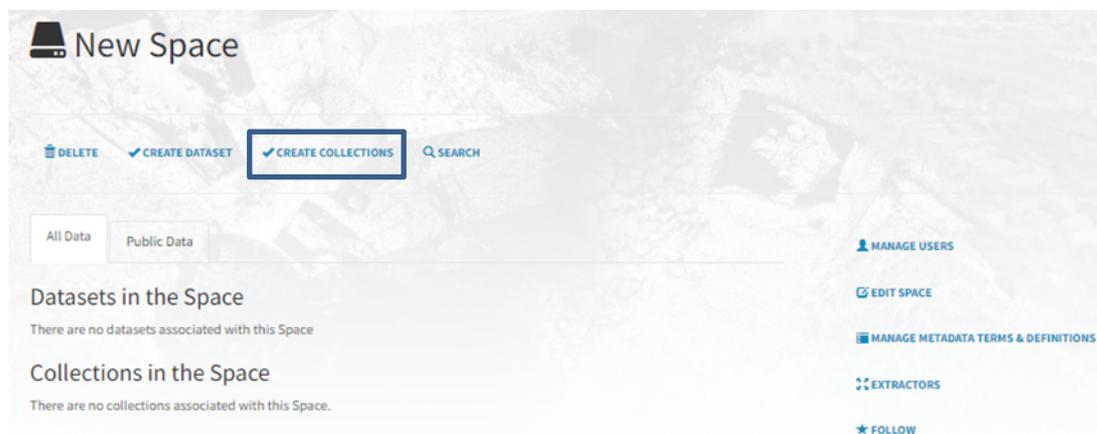


Figure 23 Creating Collections 4

In order to Create a new Collection, the only required field is the Name of the Collection. You can also include further information (Figure 24).

A screenshot of the 'Create New Collection' form. The title 'Create New Collection' is at the top. Below it is a short introductory text: 'Collections allow you to bring together multiple datasets and their associated files in an almost unlimited number of ways. Share a collection to a Space to work with your team. Permissions will be applied according to the Space's settings.' The form contains three main input areas: 'Name:' with a large text input field, 'Description:' with a larger text area, and 'Share with Spaces:' with a dropdown menu. Below the dropdown is a small explanatory text: 'By default the collection is only accessible to the user who created it. You can share this collection with members of one Space. This is optional.' At the bottom of the form are two buttons: a blue 'CREATE' button and a grey 'RESET' button.

Figure 24 Creating Collections 5



Additional Capabilities

Within a collection are available a lot of capabilities such as:

- Editing a Collection
- [Adding a Dataset](#) to a Collection (Figure 25)

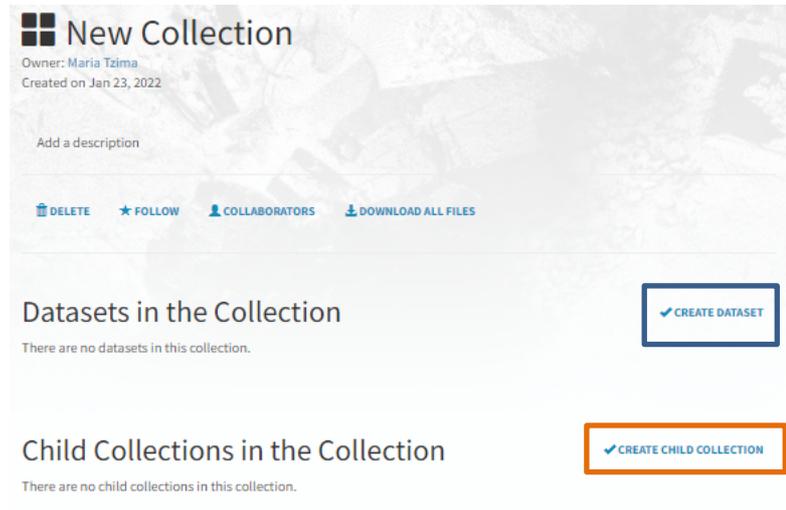


Figure 25 Additional Capabilities

- [Creating Child Collections](#)
- Removing Child Collections
- Removing from a Space



4.4 Extractors

In C4DCH, metadata can be produced by users and extractors. Users of the system can manually add metadata from the web interface by selecting from a predefined list of definitions.

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.

A wide variety of extractors have been written to support specific use cases. You can check the list of available extractors by navigating to the file within C4DCH and click on the Extractions tab.

List of Available Extractors

Audio preview	Creates image previews, thumbnails and previews of audio files
File digest	File digest extractor
Shapefiles preview	Shapefiles preview extractor takes .zip input file and communicates with GeoServer to retrieve WMS metadata. This extractor can be used to display shapefiles.
GeoTiff metadata	GeoTiff Metadata extractor, extracts metadata of a GeoTiff image. This extractor can be used in combination with the GeoTiff preview extractor.
GeoTiff preview	GeoTiff extractor communicates with GeoServer to get WMS metadata. GeoTiff preview extractor can be used to display GeoTiff files. This extractor can be used in combination with the Image preview extractor.
Video preview	Creates thumbnail and image previews of Video files, as well as downscaled video preview.
Pdf preview	Creates image previews, thumbnails and pdf previews.
NLP tika	Text extracted from the input and additional file metadata are associated with the original file.
NLP simple summary	Finds the sentence of the text which contains the most relevant words based on the text's word frequency (ignoring stop words).
NLP simple language	Detects the language of the content of a text file.
Image preview	Creates thumbnail and image previews of Image files.
Image ocr	Simple OCR (Optical Character Recognition) extractor to extract text from an image.
Image metadata	Extracts metadata from Image files.



Geospatial Previewers and Extractors

For the needs of a specific use case, extractors and previewers can work together with external services to extend the capabilities of Clowder. A typical example of this method, is the visualization of geospatial data in order to produce an interactive map.

Description:

This implementation includes a GeoTiff extractor, a shapefile extractor, a geospatial previewer and an instance Geoserver. If enabled, when a user uploads a GeoTiff or a zip file containing a shapefile, the system provides an interactive map on the dataset or file page that visualizes the geospatial raster or vector based information. Figure 26 shows how when an ESRI shapefile is uploaded to Clowder, a geospatial extractor registers the file with a Geoserver. The URL of the generated WMS layer is then added by the extractor as generic metadata back to Clowder. If generic metadata with the required keys (for example "WMS Layer URL") is available on the file, the geospatial previewer is launched and an OpenLayers-based map is rendered together with the WMS layer. It is worth noting that the previewer and extractor in this case work together but are also independent. The metadata can be added using a different extractor or client and the metadata can be used by a different previewer or client.

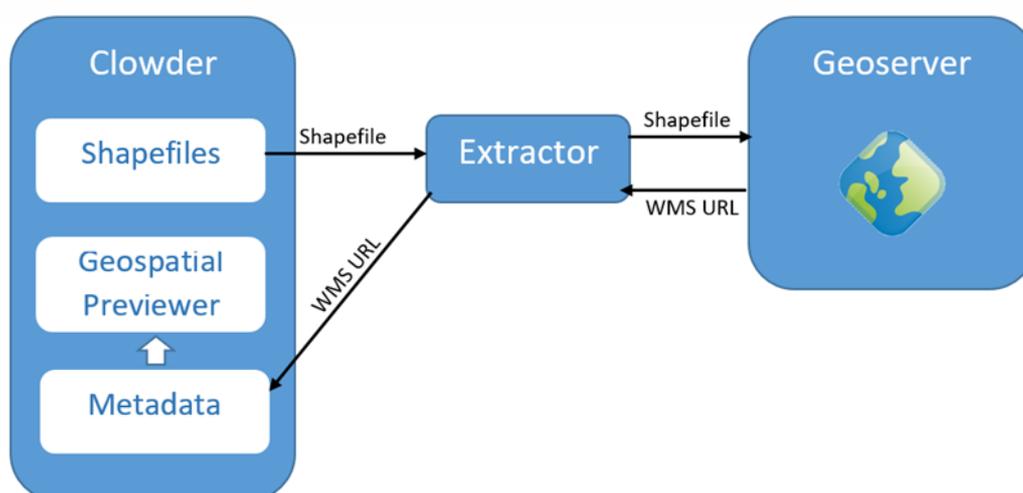


Figure 26 Geospatial support using metadata extractors and data previewers



Using Extractors

Once you have uploaded your files, you can use the available extractors according to the data type of the file. Click the Submit for Extraction Button to check the available extractors, as it is shown in Figure 27.

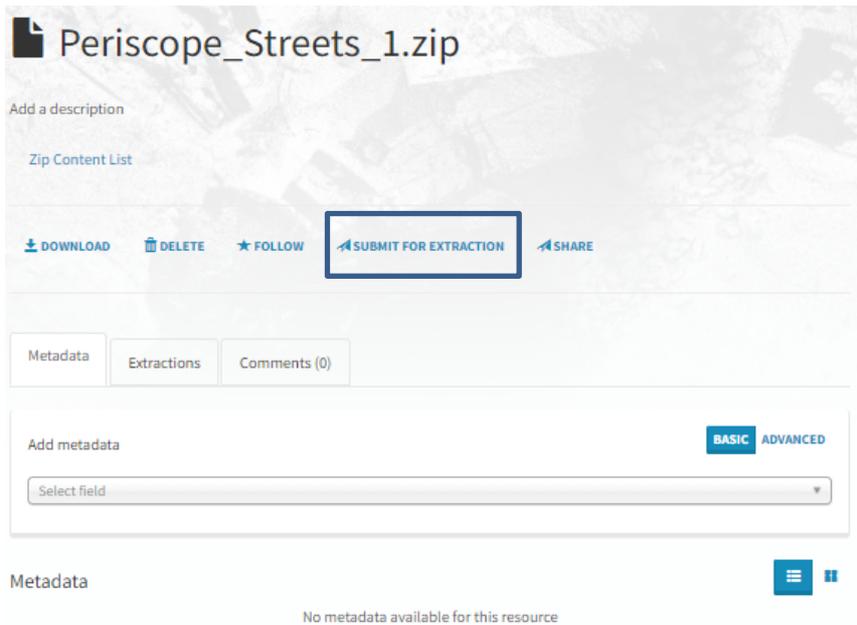


Figure 27 Using Extractors 1



Now you are able to see the list of available Extractors (Figure 28). Click the Submit Button to trigger the corresponding extractor, according to the data type of the file. For this example, the Shapefile Preview Extractor is used.

Submit file for extraction

Submit this file to a specific extractor below by providing parameters and clicking the submit button. Some parameters may be left empty.

File name: Periscope_Streets_1.zip

Extractor's Name	Description	Parameters	Submit
nlsa_audio_preview	Creates thumbnail and image previews of Audio files.		<input type="button" value="SUBMIT"/>
nlsa_file_digest	File digest extractor		<input type="button" value="SUBMIT"/>
nlsa_geoshp_preview	geoshp extractor takes .zip input file to communicate with geoserver to retrieve WMS metadata		<input type="button" value="SUBMIT"/>
nlsa_geotiff_metadata	Extractor to perform Geotiff.		<input type="button" value="SUBMIT"/>
nlsa_geotiff_preview	geotiff preview extractor takes .tif input file to communicate with GeoServer to retrieve WMS metadata		<input type="button" value="SUBMIT"/>
nlsa_image_metadata	Extracts metadata from Image files.		<input type="button" value="SUBMIT"/>
nlsa_image_ocr	Simple OCR (Optical Character Recognition) extractor to extract text from an image		<input type="button" value="SUBMIT"/>
nlsa_image_preview	Creates thumbnail and image previews of Image files.		<input type="button" value="SUBMIT"/>
nlsa_nlp_simplelanguage	Simple language extractor.		<input type="button" value="SUBMIT"/>
nlsa_nlp_simplesummary	Finds the most representative sentence in the document by term frequency.		<input type="button" value="SUBMIT"/>
nlsa_nlp_tika	Natural language extractor based on Apache Tika.		<input type="button" value="SUBMIT"/>
nlsa_pdf_preview	Creates thumbnail previews of PDF files, as well as a copy for PDF preview.		<input type="button" value="SUBMIT"/>
nlsa_video_preview	Creates thumbnail and image previews of Video files, as well as downscaled video preview		<input type="button" value="SUBMIT"/>



Figure 28 Using Extractors 2



Now, you can check the results! In this example, the .zip file containing a shapefile is visualized in an interactive map, with the use of GeoServer, as can be seen in Figure 29.

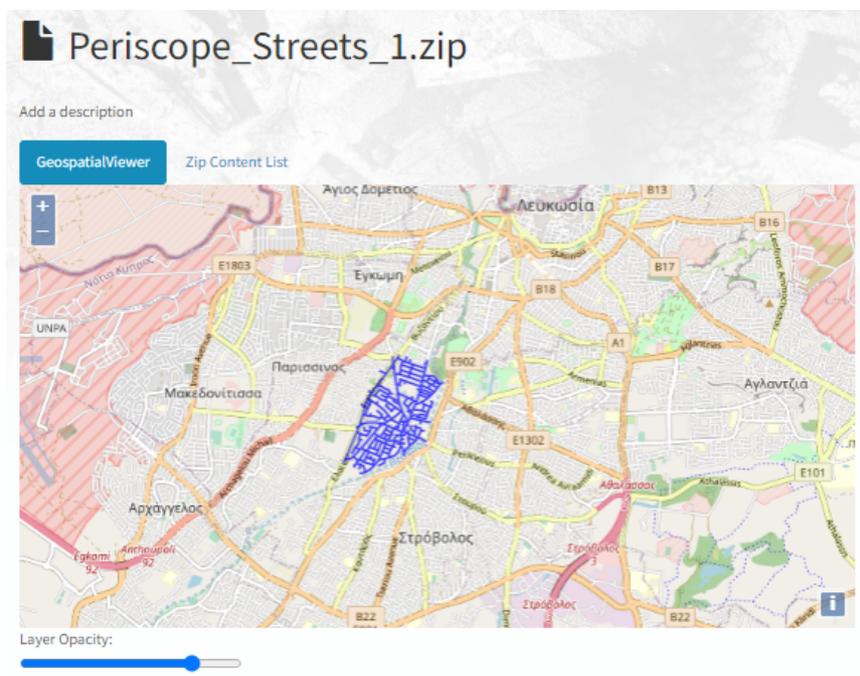


Figure 29 Shapefile preview Results



4.5 Previewers

C4DCH provides tools that are used to preview the content of large files and visualize the information contained in files and datasets in a meaningful way. These tools are called previewers and they are triggered automatically, when a file is uploaded, according to the data type of the file.

Previewer can work together with extractors and external services.

List of Available Previewers

Audio previewer	Audio previewer is a previewer for .mp3, .mpeg and mpeg3 files.
Html previewer	Html previewer is a previewer for .html files.
Plain text previewer	Shapefiles preview extractor takes .zip input file and communicates with GeoServer to retrieve WMS metadata. This extractor can be used to display shapefiles.
Viewer hop	Viewer hop is a previewer for .ply and .nxz 3D object files.
Three.js Previewer	Three.js is a previewer for .fbx 3D object files.
3D PDF	3D pdf is a previewer for .u3d files.
PDF previewer	Pdf previewer is a previewer for .pdf files.
Person tracking	Person tracking is a previewer for video files. This previewer shows you when different subjects enter the frame and leave the frame and even where in the video those subjects are.



An example of how a previewer works is displayed in the following Figure. In this example, we are uploading a .ply 3D model and the corresponding previewer is triggered. The result is displayed in the Figure 30.

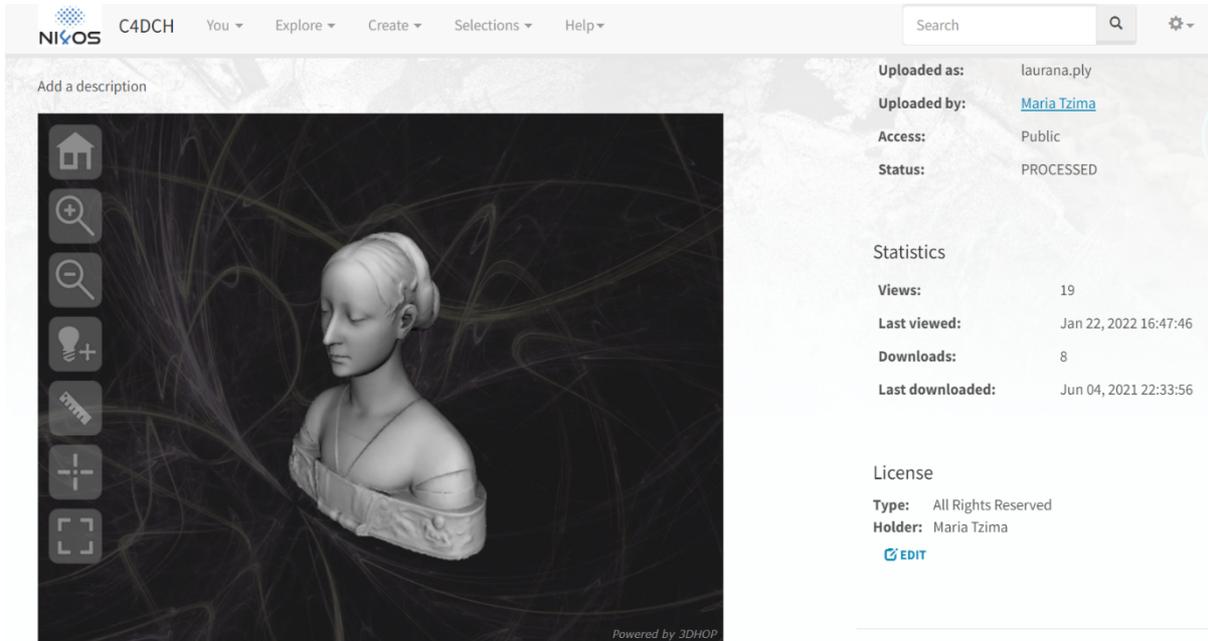


Figure 30 Previewer Example



5 Troubleshooting

