https://ada.astromat.org/?sample=OREX-803050-0

3 datasets, same title, author, description, date, different sizes and DOI

https://ada.astromat.org/doi/10.60707/pmm3-st23-- NanoIR. what is actual data; how can I figure out what they actually imaged, e.g. what 1 by 1 micron spot where on OREX-803050-0

is there a sample description for OREX-803050-0?

https://ada.astromat.org/doi/10.60707/wspc-wg10 data are tif files; UVF has "pixelSize: 0.09560" (no units), vis image has "numPixelsX: 1280

numPixelsY: 908"

The spatial resolution and image size should be apparent in the description, are the images registered-- e.g. do they overlay?

https://ada.astromat.org/doi/10.60707/and7-s828

description: "Survey data on OREX-830087-101; separate file folders for Images, Maps, and Point Spectra" what kind of images, maps, point spectra? the Files list just lists the zip archive, not the actual file content.

https://ada.astromat.org/doi/10.60707/mh28-h390

description: "This bundle includes raw images taken by the Quantitative Reflectance Imaging System (QRIS) of the sample moved from the TAGSAM to Tray A2."

? visible light? what channels are in the QRIS image, resolution? image size, how many images-- all same frame or different part of sample? Is this the entire sample acquired from the platform, or only part of it?

https://archive.astromat.org/openapi

https://archive.astromat.org/api/record/

Here’s a more data rich result, filtered to dois with metadata submitted https://archive.astromat.org/api/record/?processStatus=Calibration%20and%20Validation||Published

images--

need better classification system.

have EMPA WDS images-- depend on spectrometers used; EDS-- have orbital transition and element for energy bands, fluorescence, visible light (plane polarized, reflectance, not polarized),LIT images of various sorts?, backscatter electron, transmitted electron, electron difraction, secondary electron

are spatially registered images necessarily coupled with a basemap image.

NanoIR map collection: If maps will not be spatially registered [what is a non-spatially registered map?]: Data Collection;

If maps will be spatially registered: Spatially Registered Image Data Collection.

??? '...map' should denote spatial registration, otherwise should be '...image'.

10.60707/7479-bn03 title "MC-ICP-MS Zn Isotopes of Bennu aggregate samples OREX-803015-0 and OREX-501043-0" content is a zip archive, generalType is image. YAML for archive just says "Raw Zn Isotope Data from Neptune Plus". What is in the zip archive?? is it actually an image, or tabular data. "dataComponentType": "MCICPMSCollection"--why is it a 'collection'. is it really multicollector data, or multi sample ???

"name": "20240215\_IC\_JSC-ARES\_OREX-803001-111\_1\_ICTabular\_93.csv",

...

"generalType": "Image",

csv file is not an image....

"name": "20240215\_IC\_JSC-ARES\_OREX-803001-111\_1\_methodDescription\_410.txt",

....

"generalType": "Image",

txt file is not an image.

have to distiguish a data collection [maybe series?] (set of files with similar information content, syntax, and structure, like the set of images for an XCT) from composite dataset-- a dataset consisting of a set of distinct file types, like GIS Shape file or SLS partial scan files. The metadata needs to provide some information about the datatypes (componentType, file type...) in the collection. e.g. If there are different kinds of images (e.g. UVF and visible light), there should be a separeate image detail for each kind of image. Thus a collection might have several detail entries for images, tabular data, documents, etc included in the collection, particularly if the collection is a composite dataset.

Does anyone understand how spatial registration of points or areas on sample surfaces is working? SRS is defined in basemap image. Seems we need coordinates for fiducial marks in the basemap record, and in any spatially registered data. The basemap has to define

ada questions

supDocType = methodDescription in subjects. Where is the actual method description file??? If the description is used for many datasets, can it be a separate resource ? Or better to bundle everything in a data delivery package. ?? have table of unique files with info about file, and separate simple correlation table mapping record\_id to file\_id to associate file with a package.

.csv, ...collection.zip, methodDescription\_1.txt, methodDescription.pdf, .mp4, .emd files have general type=image???

in the instrument section, the identifier should identify an instrument instance. Might also want an identifeir for instrument type. use instrument.name as user-friendly label to identify the instrument

all contributors are 'Researcher'; are there really no other roles taht could be usefully assigned?

in 10.60707/j16b-6288, "Major element composition of olivine and carbonates measured for oxygen isotopic compositions by SIMS in OREX-800045-103" have:

{ "label": "MnO(Mass%)",

...

"fieldDescription": "Abundance of MnO in wt%" }

is it wt % or mass%?

DSC data-- have a bunch of 'other' csv files (e.g. "name": "20241111\_DSC\_NU\_OREX-803224-0\_1\_other\_2.csv","description": "Raw and Calibrated heat flow values for empty pans for step scan run 2.",

but there is no explanation of the fields in the table. All tabular data should have explanation of fields in the table...

need to represent reusable tableSchema (a la csvw) so can be called by reference, at least in a single product metadata record.

there are 161 repository itmes with the title "Analysis of particle of OREX-803077-0 by scanning electron microscopy (SEM) at Curtin University, Western Australia, including SE, BSE, and CL imaging, EDS element maps, and EBSD maps and data."

All with different DOI, and different data. This is really not very useful, and just plain lazy.

XCT data files all have the same name in record\_files table “16-bit TIFF.zip”.

Need to classify specific types for the files to one of the dataComponentTypes of SupDocTypes.

All the XanesCollections have "description": "default description",

Metadata for ‘data cubes’ doesn’t distinguish dimensions from measures.

HDF5 in pds. Use composite struture; this is an extension of what is in the guidelines.

In metadata for "specificType": "Gas Chromatography-Mass Spectrometry", "doi": "10.60707/yv1f-jb20", have 3 dimensions: but don’t know what the dimensions vs. the measure are. What are the lengths (shape) of the Time, Mass dimension (how many Times, how many ‘Mass’ s?. Have to guess tah the measure is Intensity, but what is the index order for Time and Mass dimensions? The data are in a zip archive, so can’t actually tell how its searialized—it assumes the data consumer will just figure it out?

"dimensions": [  
 {  
 "dimension": "Time",  
 "unitOfMeasure": "minutes",  
 "fieldDescription": "This dimension is the chromatographic retention time"  
 },  
 {  
 "dimension": "Mass",  
 "unitOfMeasure": "m/z",  
 "fieldDescription": "This dimension is the mass-to-charge ratio measured by the mass spectrometer"  
 },  
 {  
 "dimension": "Intensity",  
 "fieldDescription": "This is a unitless measure of intensity of the mass peak"  
 }  
 ],

For "specificType": "Transmission Electron Microscopy",  
 "doi": "10.60707/emjy-za09",

We have :

"dimensions": [  
 {  
 "dimension": "X",  
 "unitOfMeasure": "eV",  
 "fieldDescription": "Energy"  
 },  
 {  
 "dimension": "Y",  
 "unitOfMeasure": "um",  
 "fieldDescription": "Height"  
 },  
 {  
 "dimension": "Z",  
 "unitOfMeasure": "um",  
 "fieldDescription": "Width"  
 }  
 ],  
 "description": "071124\_EDS Specimen 1 Site 4 Map Data 4.h5oina",  
 "dataComponentType": "STEMEDSCube"

Here, X is apparently the measure, and Y and Z are appraently the dimensions.