

Member Nodes - Task #8830

MNDeployment # 7865 (Planning): Hydroshare (CUAHSI)

Requirements Analysis

2019-08-05 16:57 - Amy Forrester

Status:	New	Start date:	2019-08-05
Priority:	Normal	Due date:	
Assignee:	John Evans	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:			
Story Points:			
Description			
From Jeff - meeting follow-up			
links to HydroShare's site map and links to example published resources that have a level of complexity that will give you some idea of what HydroShare resources look like. Please feel free to share with your developers as I'm not sure I have their emails, and let us know if you have questions.			
Site Map:			
https://www.hydroshare.org/sitemap/			
Note that the site map includes all resources that have been shared publicly in HydroShare, not just formally published resources. Per our discussion last week, I think we would want to focus on those that have been formally published.			
Example published resources:			
https://www.hydroshare.org/resource/6711ca843bcc4834b51288a3ddc3aa08/			
https://www.hydroshare.org/resource/a3a9ba772aac4cba9533b35bb6b5fe42/			
https://www.hydroshare.org/resource/0886014659d443d4a2b9b189c96bab68/			
https://www.hydroshare.org/resource/6cd3728f69cb4cb39c6f11baac1734ec/			
If you want to access/download the BagIt bag for one of these resources, click on the link to go to the resource landing page, scroll down to the "Content" area, and click on the toolbar button that looks like a toolbox.			
All of the resources will have the same base folder structure within the BagIt bag. When you unzip a resource, you will get a folder named with the resource's GUID. Inside that is the BagIt manifest files and a folder called "data". Inside the "data" folder is the resource-level resource map (resourcemap.xml) and the resource-level Dublin core metadata file (resourcemetadata.xml). There is also a "contents" folder, which is the content folder that is displayed on the resource landing page. Users can customize the folder structure within the "contents" folder. All of the files within the "data" folder can be accessed individually through our API.			
Regarding the checksums – I spoke with one of our main developers, Scott Black, this morning. While I don't think you can currently access the BagIt manifest tag file that has the checksum information for every file in the resource through our API, he thought it would probably be pretty straightforward to expose that via our API if that is an option. Right now, that's the only place the checksums exist.			

History

#1 - 2019-08-06 14:44 - John Evans

Have gone as far as checking a single resource, the first one on the list. We may need some changes to our metadata validator, as it fails on the resourcemetadata.xml file retrieved from this "Bagit" bag (just a zip file).

Looking at crawling their sitemap to confirm the rough numbers of the published resources.