archivers.space

Information gathered during Data Rescue Events through the archivers.space app. See
http://www.ppehlab.org/datarescueworkflow#retiredworkflow for more information.

Priority *
Number should be 1-10

1

URL *

http://www.goes-r.gov/users/past-2013-01-05.html

Agency *

National Oceanic and Atmospheric Administration

Event

Title

AMS 93rd Annual Meeting - January 5-10, 2013

Crawled by the Internet Archive? *

◉ Yes

◯ No

Internet Archive URL

https://www.archivers.space/urls/3CB9C528-DFD8-4378-9E4E-F77C5D23B3D7

Description

Purpose or significance of data

Presentations from an American Meteorological Society meeting. Some of the preseos are related to GOES-R weather satellites

☐ Do not harvest. All data is small, unstructured, and on a page crawlable by the Internet Archive.

☐ Page contains dynamic content (e.g., links loaded by JavaScript).

☐ Page contains interactive visualizations.

☐ Data is accessible in structured file(s) that can be directly downloaded.

☐ Data is accessible over FTP.

☐ Data is accessible using a documented public API.

☐ Data is only accessible using search queries in a web form.

Recommended approach to harvesting data

Page contains links to pdfs of papers from the AMS Jan 2013 meeting. PDFs were harvested by looping through the HREF links with BeautifulSoup and downloading via WGet.

File formats

pdf

Estimated size in MB

396

Related URLs

http://www.goes-r.gov/users/past-2013-01-05.html

Where you able to capture all of the data at this URL?

◉ Yes

○ No

Harvest method used

PDFs were harvested by looping through the HREF links with BeautifulSoup and downloading via WGet.

Notes from Harvest

I noticed that all the downloads on this page are rooted off of http://www.goes-r.gov/downloads/. Looking at that URL - there are a lot more things in the folder besides what is referenced by this webpage. I did not download anything besides what's referenced by this page. However baggers might want to take a look at http://www.goes-r.gov/downloads/ and see if everything in there should be captured.

User certified that to the best of their knowledge this is a well-checked bag that will survive out of context of the site it was harvested from.

◉ Yes

Notes from Bagging

Bag is valid and in s3. Links to pdfs in html file are broken but remain reference-able. - aclemens

Notes from Describe

This dataset belonged to the existing CKAN record at https://www.datarefuge.org/dataset/ams-93rd-annual-meeting-january-5-10-2013, and so was added as a data resource to that record.

This form was created outside of your domain.

Google Forms