

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

3

URL *

http://satelliteconferences.noaa.gov/2013/adgen_list.htm

Agency *

National Oceanic and Atmospheric Administration

Event

Title

NOAA Satellite Conference 2013 - Presentations

Crawled by the Internet Archive? *

Yes

No

Internet Archive URL

<https://www.archivers.space/urls/83B7A19B-278E-411E-9996-0D997D09AF4C>

Description

Purpose or significance of data

This is a conference schedule with list of speakers/talks and associated video, pdfs, presentations, etc.

- 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

Grab all URLs on page (all of them are PDF, PPT, or video) run wget script.

File formats

PDF, PPT, .MOV

Estimated size in MB

3.86GB

Related URLs

http://satelliteconferences.noaa.gov/2013/adgen_list.htm

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

Yes

No

Harvest method used

wget

Notes from Harvest

wget --mirror --no-parent -r http://satelliteconferences.noaa.gov/2013/adgen_list.htm

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

Notes from Describe

This form was created outside of your domain.

Google Forms