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

10

URL *

http://nsidc.org/data

Agency *

National Aeronautics and Space Administration: Science Mission Directorate

Event

Title

Data at NSIDC | National Snow and Ice Data Center

Crawled by the Internet Archive? *

🔘 Yes

⚪ No

Internet Archive URL

https://www.archivers.space/urls/63C63237-1632-4453-8BB5-3A7927E90767

Description

Purpose or significance of data

National Snow & Ice Data Center - scientific data about the Earth's frozen regions. 1044 data sets available. Must log in (free registration) to be able to download.

☐ 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

Instructions for bulk downloads: register (free) and follow https://nsidc.org/support/faq/what-options-are-available-bulk-downloading-data-https-earthdata-login-enabled

File formats

HDF5

Estimated size in MB

Count be in the TB range

Related URLs

http://nsidc.org/data

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

🔘 Yes

⚪ No

Harvest method used

Semi-automated crawling using wget.

Notes from Harvest

I successfully downloaded all XML metadata files for 1044 databases. Further notes are in tools/README.md

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