

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

4

URL *

<http://www.eia.gov/electricity/data/eia411/>

Agency *

DOE: EIA

Event

Title

Form EIA-411 Data

Crawled by the Internet Archive? *

☒ Yes

☐ No

Internet Archive URL

https://web.archive.org/web/2019*/http://www.eia.gov/electricity/data/eia411/

Description

Purpose or significance of data

The EIA 411 report, aka "Coordinated Bulk Power Supply and Demand Program Report" collects electric reliability information from the Nation's power system planners about the electricity supply, both capacity and energy, that is needed to serve current demand and for future growth.

The source site provides data from 1990-2014. The source says the last update date is May 16th 2016, and the next data release date is November 2016, but as of February 2017, it still has yet to be updated.

- ☐ 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

All of the relevant links can be pulled from here:

http://www.eia.gov/electricity/data/eia411/eia411_history.html.

We only need to pull in the XLS files. The PDFs are duplicates of the same data. Note: each of the XLS files contains multiples tabs of data.

File formats

XLS

Estimated size in MB

Related URLs

www.eia.gov/electricity/data/eia411

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

☒ Yes

☐ No

Harvest method used

Wrote a python script using BeautifulSoup to automatically download all .xls files. Grabbed 3 images manually.

Notes from Harvest

Original researcher pointed to "http://www.eia.gov/electricity/data/eia411/eia411_history.html" but that should technically be a different url project as this page's own excel files were neglected in the original write-up. The history url will be added as a link to this url.

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