

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

5

URL *

http://afdc.energy.gov/vehicles/electric_emissions.php

Agency *

United States Department of Energy

Event

Title

Emissions from Hybrid and Plug-In Electric Vehicles

Crawled by the Internet Archive? *

Yes

No

Internet Archive URL

https://web.archive.org/web/*/http://afdc.energy.gov/vehicles/electric_emissions.php

Description

Overview of tail pipe emissions from hybrid vehicles; comparisons to other vehicles

Purpose or significance of data

- 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

Python scripts to submit query data to download the Highchart maps.

File formats

svg

Estimated size in MB

7.5

Related URLs

afdc.energy.gov/vehicles/electric_emissions.php

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

Yes

No

Harvest method used

Captured data from API (emissions by state) and saved locally to make interactive visualization functional. Also saved a few core webpages and PDFs.

Notes from Harvest

[2/26/17 1:08AM PST]: Original harvester here. There are no SVGs because researcher recommended saving each view of the interactive visualization. I was able to preserve the entire interactive viz instead, so I don't see the reason to create SVGs. So I believe the researcher wanted images of each view; I preserved the entire interactive visual instead.

[2/25/17 10:28PM PST]: Sent back up from bagging. No svg files were present inside the zip, though this seems to be the main file type of interest according to the research section.

Primary focus was making visualization work because Archive.org has website archived, but visualization is non-functional. I saved a few core webpages and PDFs related to assumptions, related reports but did not capture all links.

** NO HARVEST URL PLEASE RE-HARVEST ** UUID should show up in <https://drp-upload.s3.amazonaws.com/> if it's uploaded properly.

EDIT from Harvester: I believe UUID does show up in above link.

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

Bagged with Data Conservancy Packaging tool; validated before S3 upload and after S3 download.

Notes from Describe

Described by Erin Antognoli

This form was created outside of your domain.

Google Forms