1. Priority *

| 10 |

2. URL *

| https://data.energystar.gov |

3. Agency *

| N/A |

4. Event

| |

5. Title

| ENERGY STAR Certified Products | EPA ENERGY STAR | ENERGY STAR |

6. Crawled by the Internet Archive *

( ) Yes

7. Internet Archive URL

https://web.archive.org/web/*/http://epa.gov/dwssa

8. Description

9. Purpose or significance of data

Contains data on the features of all products that currently qualify for EPA's ENERGY STAR label.

10.

- [ ] Do not harvest. All data is small, unstructured, and on a page crawlable by the Internet Archive.
- [x] 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.
- [x] Data is accessible using a documented public API.
- [ ] Data is only accessible using search queries in a web form.

11. Recommended approach to harvesting data

via existing API at https://data.energystar.gov/developers

12. File formats

JSON

13. Estimated size in MB

14. Related URLs

https://data.energystar.gov

15. Were you able to capture all of the data at this URL?

◉ Yes

◯ No

16. Harvest method used

Scraping JSON from public API

17. Notes from Harvest

Three references were broken: ikxg-vjvt.json (energy efficient servers, found and saved at: yhae-jcp4.json ij6p-f5pr.json They were manually downloaded via the front-end interface and paired with the blank json from the three broken API links. I manually downloaded their views and included them in the dataset.

18. 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

19. Notes from Bagging

Bagged with Data Conservancy Packaging Tool (v.1.0.5); validated before S3 upload and after S3 download.

20. Notes from Describe