4 - Anonymous ∨

05:01

Time to complete

< >

1. Priority *

10

2. URL *

https://echo.epa.gov

3. Agency *

N/A

4. Event

5. Title

Enforcement and Compliance History Online

6. Crawled by the Internet Archive *

○ Yes

7. Internet Archive URL

https://web.archive.org/web/*/https://echo.epa.gov

8. Description

9. Purpose or significance of data

Use EPA's Enforcement and Compliance History Online website to search for facilities in your community to assess their compliance with environmental regulations. You can also investigate pollution sources, examine and create enforcement-related maps, or explore your state's performance. Learn more about ECHO

10.

☐ 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.

11. Recommended approach to harvesting data

Data can be obtained via: Direct download: https://echo.epa.gov/tools/data-downloads#downloads REST services: https://echo.epa.gov/tools/web-services

12. File formats

ZIP, PDF

13. Estimated size in MB

14. Related URLs

echo.epa.gov

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

◉ Yes

◯ No

16. Harvest method used

link scraper, wget loop

17. Notes from Harvest

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