

Responder 29



15 - Anonymous



02:37



Time to complete

1. Priority *

10

2. URL *

<http://epa.gov/waterdata/waters-watershed-assessment-tracking-environmental-results-system>

3. Agency *

N/A

4. Event

5. Title

WATERS Mapping Services

6. Crawled by the Internet Archive *

☒ Yes

7. Internet Archive URL

https://web.archive.org/web/*/http://epa.gov/waterdata/waters-watershed-assessment-tracking-environmental-results-system

8. Description

9. Purpose or significance of data

water quality data - data found in downloadable files on the Get Data/Tool tab. May be useful for climate studies if temperature is recorded.

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

12. File formats

13. Estimated size in MB

14. Related URLs

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

☒ Yes

☐ No

16. Harvest method used

Link scraper/wget

17. Notes from Harvest

I didn't download any of the data that came from completely separate subdomains

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

About 3.6 GB

20. Notes from Describe

3.6 GB