

Responder 29



29 - Anonymous



02:37

Time to complete



1. Priority \*

8

2. URL \*

<https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=308341>

3. Agency \*

N/A

4. Event

5. Title

EPA's Study of Hydraulic Fracturing and Its Potential Impact on Drinking Water Resources: Compilation of Physicochemical and Toxicological Information About Hydraulic Fracturing-Related Chemicals (Draft Database)

## 6. Crawled by the Internet Archive \*

Yes

## 7. Internet Archive URL

[https://web.archive.org/web/\\*/https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=308341](https://web.archive.org/web/*/https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=308341)

## 8. Description

## 9. Purpose or significance of data

This product is being released in support of the hydraulic fracturing study (HF Study) assessment and is available to provide the information that was used to determine the potential impacts of hydraulic fracturing for oil and gas on drinking water resources in the United States. This is companion data to the draft report that was released for public comment and shared with the EPA Science Advisory Board for expert peer review. [Press Release Jun 4, 2015] The purpose of this product is to make accessible the information about the 1,173 hydraulic fracturing-related chemicals that were listed in the external review draft of the Hydraulic Fracturing Drinking Water Assessment that was released recently. The product consists of a series of spreadsheets with physicochemical and toxicological information pulled from several sources of information, including: EPI Suite, LeadScope, QikiProp, Reaxys, IRIS, PPRTV, ATSDR, among other sources. The spreadsheets also contain background information about how the list of chemicals were compiled, what the different sources of chemical information are, and definitions and descriptions of the values presented. This is a subsection of the website for "hfstudy" titled: EPA's Study of Hydraulic Fracturing for Oil and Gas and Its Potential Impact on Drinking Water Resources

## 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

Click "Downloads" and save zip files. There are 10 .xls files.

#### 12. File formats

ZIP, XLSX

#### 13. Estimated size in MB

4.3

#### 14. Related URLs

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

Yes

No

#### 16. Harvest method used

Manual download

#### 17. Notes from Harvest

Downloaded zipx file and packaged -- I manually downloaded all PDFs and xls files I was able to find at the study pages. All HTML pages should have been captured by the wayback machine. The research notes suggest that there are only xls files, but I also found pdfs associated with the study itself, and included those with the data.

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

Got the harvest url. Contains XLS files.

#### 20. Notes from Describe