

# 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

3

---

## URL \*

<http://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/mrip-project>

---

## Agency \*

National Oceanic and Atmospheric Administration

---

## Event

---

## Title

MRIP Projects

---

## Crawled by the Internet Archive? \*

☒ Yes

☐ No

## Internet Archive URL

<https://www.archivers.space/urls/ABE6552C-A596-41F3-94F6-EA91743283D0>

---

## Description

---

## Purpose or significance of data

### Projects and PDFs

Since 2008, MRIP has funded survey improvement projects and survey program reviews in all regions of the U.S. and its Territories. Projects span everything from utilizing electronic reporting tools to mail surveys to increasing precision of estimates. On this page, you can find all funded projects since 2008, read project plans, final reports, and monthly updates.

---

- ☐ 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

loop through iframed data content panes and download pdfs for each project

---

## File formats

PDF

---

## Estimated size in MB

200

---

## Related URLs

<http://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/mrip-project>

---

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

☒ Yes

☐ No

## Harvest method used

python selenium interacting with iframed database to download pdfs

---

## Notes from Harvest

all went well, I downloaded each project into a separate folder, within which is the page source and downloaded pdfs. The page source has the description in the database

---

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

---

Notes from Describe

---

This form was created outside of your domain.

Google Forms