

# 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

N/A

---

## URL \*

<http://marineprotectedareas.noaa.gov/dataanalysis/mpainventory/>

---

## Agency \*

N/A

---

## Event

N/A

---

## Title

Marine Protected Areas Inventory Data

---

## Crawled by the Internet Archive? \*

☒ Yes

☐ No

## Internet Archive URL

<https://www.archivers.space/urls/06706FA3-EFBD-43C1-844B-993271F1F204>

---

## Description

---

## Purpose or significance of data

---

- ☐ 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

Download tabular data and GIS data

---

## File formats

shp, xls

---

## Estimated size in MB

---

## Related URLs

<http://marineprotectedareas.noaa.gov/dataanalysis/mpainventory/>

---

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

☒ Yes

☐ No

## Harvest method used

Just downloaded the data from the site directly

---

## Notes from Harvest

Easy one.

---

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

Bagged with Data Conservancy Packaging Tool; validated before S3 Upload + After S3 Download

---

# Notes from Describe

Tagged in CKAN with "Sustainable Development Goals" and marked in bold at bottom of description with the specific indicator (Sustainable Development Goal Indicator 14.5.1 - Coverage of protected areas in relation to marine areas)

.....

This form was created outside of your domain.

Google Forms