Simple Data Catalog

2018 April
Environmental Data Initiative (EDI)

Introduction

Five Phases of data publishing:

1. Organize data and metadata
2. Clean and format data tables
3. Create EML metadata
4. Upload data package to EDI repository
5. Cite and link

https://environmentaldatainitiative.org/resources/five-phases-of-data-publishing/

Example

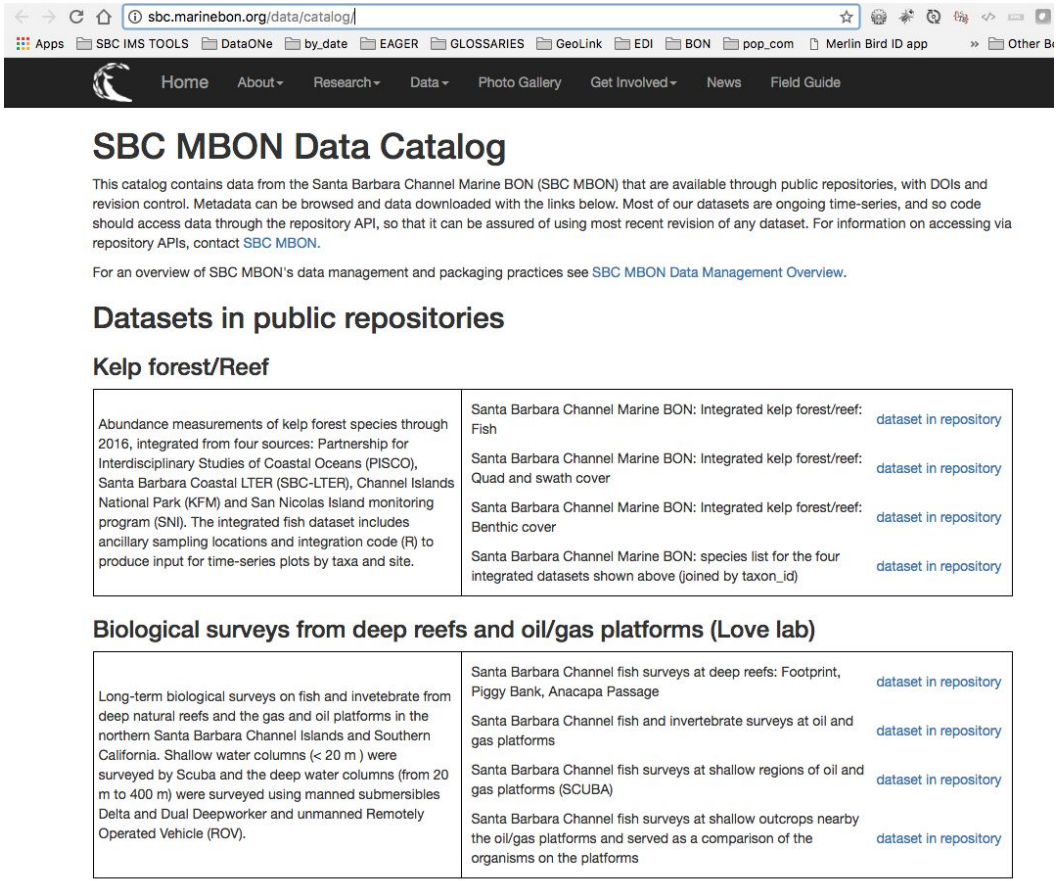
http://sbc.marinebon.org/data/catalog/

HTML has

- Text description, dataset titles
- Hyperlink to the dataset *in the EDI repository*

Good solution for groups that

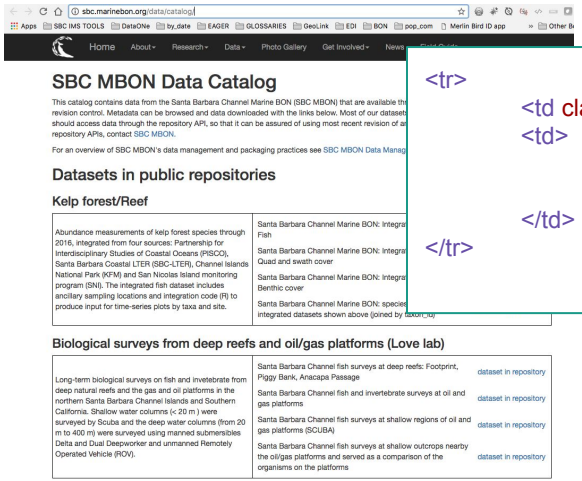
- Have few datasets
- Don't need a custom display
- Don't have time to code



Using the identifier

Example: hyperlink to the most recent revision of dataset “edi.5”:
https://portal.edirepository.org/nis/mapbrowse?scope=edi&identifier=5&revision=newest

dataset in repository



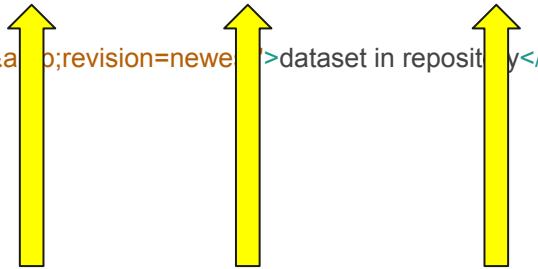

```
<tr>
        <td class="ds_title">CALCOFI fish larvae at 66 standard stations, 1966 - ongoing</td>
        <td>
                <a href="https://portal.edirepository.org/nis/mapbrowse?scope=edi&amp;identifier=109&amp;revision=newest">
                dataset in repository</a>
        </td>
</tr>
```

Using the identifier

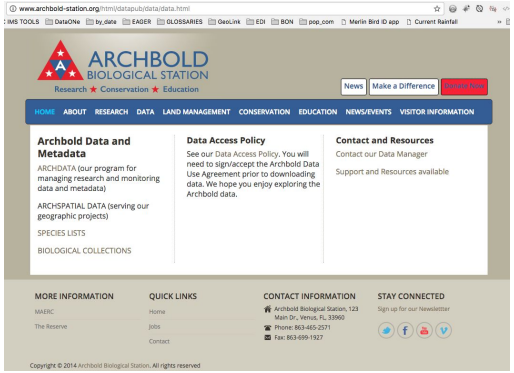
Example: hyperlink to the most recent revision of dataset “edi.5”:

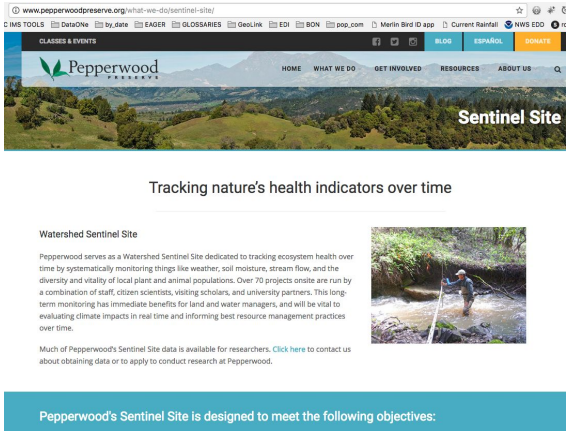
https://portal.edirepository.org/nis/mapbrowse?scope=edi&identifier=5&revision=newest

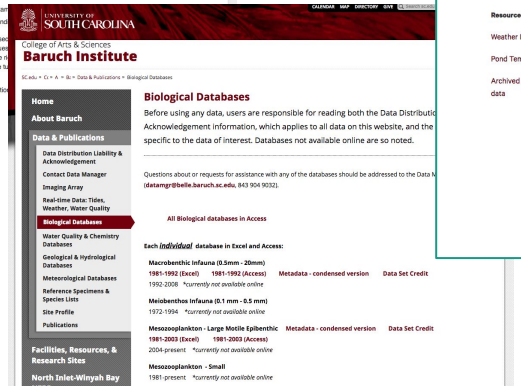
dataset in repository

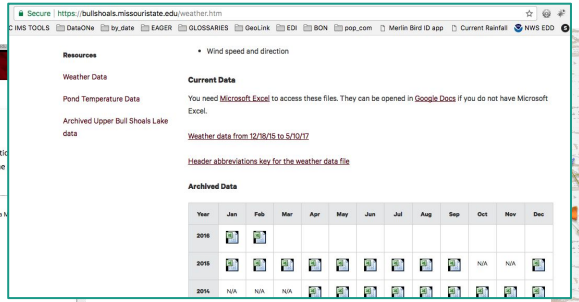


Add to an existing page …









Dataset links

DOI

Starts with “10. “

Permanent, to a specific revision:

https://dx.doi.org/10.6073/pasta/62803c95783c4e771695d1c6cc3d23ac

Repository identifier

Most recent revision:

https://portal.edirepository.org/nis/mapbrowse?scope=edi&identifier=5&revision=newest

Which link?

To	Use	Result
Cite data...	DOI	Resolves to the exact revision
Link to an ongoing time series, where you plan to update with new revisions...	Use the repository id, with revision=”newest”	User will be taken to the most recent revision of your time series User will see the DOI, so it can be cited correctly

Duane Costa

Complex catalogs with PASTA API