

Gulf of Maine Ocean Data Partnership – Just how accessible is our data?

June 26, 2007

Background

In January 2006, Evan Richert, member of the Executive Committee, posed the question of how accessible the partnership’s data were. He also wanted to gain a better understanding of what “accessibility” means. To answer these questions, a subcommittee of the Technical Committee was formed consisting of Bob Branton, Bob Groman, Lou Van Guelpen, Jim Manning, Sylvia Most, Joan Palmer, Deb Soule, and Nick Wolff.

Approach

A list of the partner’s data web links was compiled from the GCMD metadata records and through the previous surveys if a metadata record was not available. Several records were initially reviewed against draft criteria. After reviewing the initial results, the criteria were revised and the rest of the links were reviewed. (Final criteria are listed in the table below. Bolded values were anticipated reviewer answers although other answers (particularly “don’t know”) were also given.) Each subcommittee member reviewed several links but not their own. Members were not expected to make an exhaustive effort to find the data for a particular data set or to contact the data provider if questions arose. Rather they were expected to make a more cursory check to see if the data could easily be found using the link provided. The links previously reviewed under the old criteria were reevaluated based on the new criteria. The answers were compiled in worksheets in a single spreadsheet located at: <http://www.gomodp.org/technical-committee/gomodp-accessibility-survey-compiled-results.xls/download>.

Criteria	Description
Data accessible	<p>Is the data accessible via the web?</p> <p>Yes - Data are accessible if actual information can be obtained either by query, direct access or automatic email with attached results.</p> <p>No - No data could be found, link was temporarily inaccessible, human intervention was needed or was proprietary data.</p> <p>Broken link – Link provided resulted in a web error.</p>
Software needed	<p>Is special software needed to access data?</p> <p>Yes - Special client download software required.</p> <p>No</p>
Registration required	<p>Is registration required to access data?</p> <p>Yes - Personal information for purposes of authentication is required.</p> <p>No (Email address for delivery only is not considered registration)</p>
Formats	<p>What data formats are available for download?</p> <p>ASCII, NetCDF, html, spreadsheet, db, shapefile, other (describe)</p>
Protocols	<p>What data protocols are available?</p> <p>WFS, OPeNDAP, WMS, other (describe)</p>
Map interface	<p>Is there a mapping interface available for the data?</p> <p>Yes</p> <p>No</p>
Description	<p>General notes about the web access link.</p>

Findings

Statistics shown below were derived from the aforementioned spreadsheet. In some cases, answers were consolidated to fit a broader category answer. In general, about 50% of the partnership’s data links are

considered reasonably accessible. Over 80% of the data access does not require registration or the installation of software. However, many of the links could not be fully evaluated due to broken links or the reviewer's inexperience/inability to evaluate a data link for a specific category. Specific numbers are below.

Statistic Description	# of links	Percentage
Overall Statistics		
Links reviewed	183	100%
Links obtained from metadata records	173	95%
Links found only through survey (21 found but only 10 evaluated)	10	5%
Data accessibility		
Data were considered accessible	87	48%
Data were not considered accessible	88	48%
Broken links	6	3%
Accessibility could not be determined	2	1%
Software needed		
Software is needed to access data	1	<1%
Software is not needed to access data	153	84%
Don't know/couldn't evaluate if software is needed or not	29	16%
Registration required		
Registration is required before data can be obtained	6	3%
No registration is required to obtain data	150	82%
Don't know/couldn't evaluate if registration is required.	27	15%
Formats available (*Total numbers will exceed 183 since data sometimes can be saved in many formats from the same link.)		
ASCII, Text	45	22%
Don't know/couldn't evaluate	95	46%
ESRI GIS/Shapefile/ArcInfo	27	13%
Excel	5	2%
HDF	2	1%
HTML	26	12%
NetCDF	3	1%
Other (MapGen, MatLab, , PDF, Splus, Tiff, WMS – 1 each)	6	3%
Protocols available (*Total numbers will exceed 183 since data sometimes can be accessed by more than 1 protocol.)		
Don't know/couldn't evaluate	29	15%
HTTP	3	2%
None	27	14%
OpenDAP	7	4%
Other (FTP, IMS, WMS – 1 each)	3	2%
Blank (no response listed in spreadsheet by reviewer)	124	64%
Map interface		
Yes	15	8%
No	31	17%
Don't know/couldn't evaluate	13	7%
Blank (no response listed in spreadsheet by reviewer)	124	68%

Recommendations

Based on this exercise, the subcommittee has the following recommendations:

- Inform individual providers of broken links.
- Share information gleaned with the rest of the partnership.

- Update technical guidance to request the use of “GetData” links in the metadata records. Also, specify that link should go to where the data resides – not at the organizational or some other upper level.
- In order to facilitate an easier accessibility review in the future, increase partnership’s use of metadata records and appropriate “GetData” or “Related” links.
- Some provider records could be simplified such as not separating a data set into metadata records for each year of data.
- Ask Melanie Meaux to create a master metadata record in the GCMD GoMODP portal that is a list of partnership entries with the date each record was last updated. This would aid in future reviews of any type.
- For any future reviews, provide greater guidance to reviewers on how to evaluate the links consistently, accurately, and confidently.