LTER

01001100 01010100 01000101 01010010

LTER DataBits

Information Management Newsletter of The Long Term Ecological Research Network

01001100 01010100 01000101 01010010

Web-Based Data Visualization With JPGraph

- Mason Kortz (PAL and CCE)

With the increasing size and complexity of datasets available on the web, in-place data visualization is becoming more important. While most web-based visualization tools lack the capacity for actual data analysis, they are still very useful for finding datasets of interest, previewing data, and performing quality control. One tool for web-based data visualization is JPGraph, a PHP plotting library available as both a free download and a paid, professional version. The professional version is licensed for commercial applications and includes more support, bar code functionality, and windrose and odometer graph types. This article discusses the benefits and drawbacks of the free version of JPGraph.

JPGraph is a library of PHP classes that can be used to create many types of graphs, including line, bar, scatter, and error plots. Versions are available for both PHP 4 and PHP 5. PHP must be compiled with GD support enabled to support JPGraph. Further, in order for JPGraph to access TrueType fonts, PHP must be compiled with TrueType font support. This fonts feature allows the use of superscripts, subscripts, and special characters such as symbols and Greek letters. Information on both these compilation options can be found in the PHP manual.

One of the advantages of JPGraph is that, being a PHP library, all execution is done on the server. Unlike Java, Javascript, or Flash visualization tools, the output of JPGraph is not dependent on the client's software or configuration. JPGraph's class structure is powerful and grants a great deal of control over the content and appearance of graphs. Callback functions can be applied to all axes, allowing user-defined manipulation of data points where the built-in functions are not sufficient. JPGraph allows multiple Y-axes with independent scales, and allows multiple plots to be overlaid on a single graph, allowing for graphs such as line plots with error bars or scatter plots in multiple colors. JPGraph can handle fairly large datasets; graphs with over 10,000 data points can be plotted in just a few seconds. Finally, and perhaps most importantly, JPGraph's class structure is very well documented and numerous examples are available both in the manual and on the website.

JPGraph does have certain limitations. Because the output of JPGraph is an image, it does not have the capability for interaction that Java- or Flash-based graphs may have. Graphs can be associated with image maps, but this becomes impractical with even moderately large datasets. Date and time axes are handled as UNIX timestamps, which means dates prior to January 1st, 1970 must be handled via a user-defined callback function in order to plot correctly. Similarly, inverting axes requires a callback function. While these functions are not difficult to define, they do add to the overhead of displaying graphs with many data points. There is no support for graphs with multiple X axes. In addition getting the proper positioning for labels, legends, and titles is often a matter of trial and error. In terms of dataset size, JPGraph has it's limits - graphs with over 100,000 data points can take over a minute and may cause the client browser to time out before completion.

Overall, JPGraph provides a free, straight-forward, and well-documented answer to online plotting. While it lacks the power of offline, commercial solutions, it is ideal for providing a quick visual window into a dataset without a great deal of setup overhead.