



BlogMyData: A Virtual Research Environment for collaborative visualization of environmental data

Andrew Milsted (1), Jeremy Frey (1), Jon Blower (2), and Adit Santokhee (2)

(1) School of Chemistry, University of Southampton, United Kingdom, (2) Reading e-Science Centre, Environmental Systems Science Centre, University of Reading, United Kingdom

1 Introduction

In the area of Earth system modeling collaborations take place chiefly through face-to-face meetings, the scholarly literature and informal electronic exchanges of emails and documents. All of these methods suffer from serious deficiencies that hamper effective collaboration. For practical reasons, face-to-face meetings can be held only infrequently. The scholarly literature does not yet adequately link scientific results to the source data and thought processes that yielded them, and additionally suffers from a very slow turnaround time. Informal exchanges of electronic information commonly lose vital context; for example, scientists typically exchange static visualizations of data (as GIFs or PostScript plots for example), but the recipient cannot easily access the data behind the visualization, or customize the visualization in any way. Emails are rarely published or preserved adequately for future use. The recent adoption of “off the shelf” Wikis and basic blogs has addressed some of these issues, but does not usually address specific scientific needs or enable the interactive visualization of data.

2 The Solution

A Virtual Research Environment is an attractive solution to the above problems. In the JISC-sponsored BlogMyData project we are creating such a VRE by combining the capabilities of two existing technologies that have already seen wide adoption among scientists:

1. The Godiva2 data visualization system (<http://www.reading.ac.uk/godiva2>) provides a means for scientists to browse interactively in a “Google Maps-like” fashion through large environmental datasets, including numerical model outputs and high-resolution satellite imagery, using only a web browser. Scientists can produce maps, timeseries and other plot types.
2. The LabTrove is a web-based blogging tool specifically designed for the practising scientist to record, disseminate and evaluate their research. The Blog can also be used as a collaboration tool that allows discussion between colleagues. For open science work the blog uses standard protocols (such as Really Simple Syndication, RSS) to publish its content to the public domain but also contains the necessary access control to keep any private work secure. Although initially designed for the use of laboratory chemists, the LabTrove is being adapted in this project to meet the needs of environmental scientists.

Having logged in to the BlogMyData VRE using OpenID, users examine output from the latest cutting-edge climate and ocean models using Godiva2. Upon finding a feature of interest the user creates a new blog entry that is linked to the current visualization. The blog entry is automatically tagged with metadata about the feature of interest (e.g. its location in time and space, and the dataset from which it is derived). Colleagues provide input through comments and by linking blog entries together. Through semantic and geospatial tagging, scientists can discover colleagues working on similar scientific problems. The system is augmented by the addition of a geospatially-enabled database, based on the widely-used open-source PostgreSQL database with the PostGIS extensions. This database will associate blog entries with geographical areas and time periods and allow users to discover discussions that relate to particular areas of interest very efficiently.

The system is being developed with regular feedback from users in NCAS-Climate and NCEO. We have created an end-to-end prototype of the system, in which users can create blog entries based upon map-based visualizations.

Blog entries are captured in a private blog, which is only visible to a controlled set of users (authenticated using OpenID), thereby maintaining the privacy of the research. Blog entries can be syndicated as GeoRSS feeds. These feeds can be consumed in standard RSS viewers (such as Microsoft Outlook, Google Reader and Firefox Live Bookmarks), or in “geo-enabled” viewers such as Google Maps.