

Web 2.0 and the the ever elusive balance between Information explosion and data mining
Abhay Krishna, Integrative Ecology Group, Estacion Biologica de Doñana,CSIC, Sevilla , Spain

In a fascinating tussle of perspectives in Volume 6 and issue no 3 of april 2008 Frontiers, Martin A Nuñez and Gregory M Crustinger enlist ways to brace up with the recent literature. They advocate that PhD students should choose a well studied system to have better chances of success. On the other side of this fascinating tussle Daniel Simberloff and Nathan J Sanders agree with there enthusiastic case and approaches for bracing up with the recent literature. However, they disagree with the play safe strategy of start-ups in ecology and evolution.

For me this tussle of perspectives is really fascinating. It underlines the story of our times. In the times of information explosion, should we concentrate on mining this information for knowledge or should we keep adding to this information explosion. This dilemma fits well in the context of the tussle by a quote from Nassim Nicholas Taleb “ It is almost impossible these days to finish PhD without excessive intellectual curiosity and it is impossible to get a faculty position without narrowly specializing in a chosen field ”.

Here in this letter, I want to harmonize the two perspectives by arguing that coexistence of information explosion and data mining is possible. We should strive for that ever elusive balance and I discuss below how Web 2.0 will enable it. Web 2.0, still in its infancy already shows promising results and I discuss below these usages.

Web 2.0 is a term describing the trend in use of world wide web technology and web design that aims to enhance creativity, information sharing and collaboration among users. Here I describe three Web 2.0 applications having huge potential for facilitating the ever elusive balance:

1. RSS feeds

Most of the scientific journal and science related websites now offer RSS feeds which are just the log of any changes made, i.e. for scientific journals any new articles. Researchers could subscribe to the feeds of these journals and read subscribed feeds on a RSS reader for eg Google provides a web based RSS feed reader.

<http://reader.google.com>

2. Grease monkey

It is a Firefox extension that allows small pieces of JavaScript to modify the look and function of a web page. It's easy to install and sit unobtrusively down in the browser status bar, from where you can enable/disable it or add and manage scripts. Grease monkey is available at:

<https://addons.mozilla.org/en-US/firefox/addon/748>

The potential is endless and only bound by our imagination. Some examples are:

- pubmed2connotea: This script adds icons to a PubMed results page to bookmark references at Connotea, CiteUlike or del.icio.us, which are social bookmarking sites. The script is available at <http://www.urbigene.com/pubmed2connotea/>
- postgenomic script: This script adds marks articles at scientific journal pages that are discussed at Postgenomic (<http://www.postgenomic.com/>). The script is available at: http://wiki.nodalpoint.org/projects_postgenomic

Javascript and greasemonkey can be learnt by experimenting with numerous scripts available at:
<http://userscripts.org/>

3. Social networking: Web 2.0 offers endless possibilities to share, transform and collaborate on scientific ideas and projects. Few examples are:
 - Social bookmarking websites: Connotea is an excellent example, where science related links are shared and managed on web, Connotea is available at:
<http://www.connotea.org/>
 - Web 2.0 journals: Few examples are, PLOS one and Nature Preceedings, This is a refreshingly new concept, in which scientific community discuss a paper as a community thus offering insights and criticisms.
<http://www.plosone.org/home.action>
<http://precedings.nature.com/>
 - Scientific social networking: Nature networks is an example of such a community where researchers can form networks and collaborate.
<http://network.nature.com/>

This is just a glimpse of the potential of Web 2.0 . Web 2.0 can go a great distance in achieving the balance between information explosion and data mining.