Browser Based Applications Supported by R in Pipeline Pilot

DaveNicolaides[1,*], Noj Malcolm[1], Stephane Vellay[1], Dana Honeycutt[2], Tim Moran[2]

1. Accelrys Ltd., 334 Cambridge Science Park, Milton Road, Cambridge UK, CB4 0WN
2. Accelrys Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121, USA
*Contact author: dave@accelrys.com

Examples of the use of R as the "statistical engine" behind a simplified interface designed to be used by scientists abound (Bajuk-Yorgan, Kaluzny, 2010; Weiss 2008; Neuwirth 2008; there are perhaps another dozen examples in the useR! meetings over the last two years alone).

We wish to share the success we've had in deploying R more widely to the research community in standard browsers, via our Pipeline Pilot informatics platform (Hassan et al. 2006). While there are similarities between our approach and those referred to previously, we think that a reliance on the scientific domain knowledge increasingly to be found in R packages distinguishes our approach. We will give specific examples of the use of R behind the scenes, in such diverse areas of research as pharmaceutical discovery, automated image classification, gene expression, heterogeneous catalysis, and process analytical technology.

The success of these activities have largely stemmed from their approach to integration not as an activity based on software technologies, but instead based on people and the ways they work and learn. Allowing scientists to leverage a sophisticated statistics engine in a familiar interface gives them greater insight into and confidence in their decisions. In seeking to extend this success we will have to address many complex issues, including:

- distinguishing transfer of statistical expertise (where the goal is learning) from mere automation (where the goal is research efficiency), and
- understanding how we might "enable ideas as software" in R, where those ideas already have a strong presence in our own and other commercial software products.

References

Lou Bajuk-Yorgan, Stephen Kaluzny (2010). Making R accessible to Business Analysts with TIBCO Spotfire. In *useR! 2010, The R User Conference (Gaithersburg, Maryland)*, Book of Abstracts p. 14.

ChristianWeiss (2008). Commercial meets Open Source - Tuning STATISTICA with R.In *useR! 2008, The R User Conference (Dortmund, Germany)*, Book of Abstracts p. 188.

Erich Neuwirth (2008). R meets the Workplace - Embedding R into Excel and making it more accessible.In *useR! 2008, The R User Conference (Dortmund, Germany)*, Book of Abstracts p. 135.

Hassan M, Brown RD, Varma-O'brien S and Rogers D. (2006). Cheminformatics analysis and learning in a data pipelining environment. *Moleculardiversity* 10(3), 283-99.