Deployment of validated statistical algorithms within an end-user's application

Dr. Mark Demesmaeker, Spotfire Inc., www.spotfire.com

Modern life science research data demands sophisticated analysis. New technologies and new algorithmic developments require close collaboration between research- and statistical organizations. While once it was possible for statisticians and analytic experts to work one-on-one with researchers and their data sets, the pace of research now demands more rapid and broader deployment of analysis software to share analytic expertise across the organization. Increasingly statistical staffs are relying on their ability to influence the software used by researchers as their means of ensuring rigorous validation of experimental findings. Unfortunately, off-the-shelf software rarely contains the statistician's algorithm or statistical method of choice. Statisticians find themselves becoming application developers in order to provide the necessary analytics to their end-user colleagues. End-users struggle with the complexity of analysis applications, used only occasionally but at critical junctures, to validate their findings. This misalignment of resources can decrease the effectiveness of organizations better prepared to derive algorithms than developing and maintaining end-user software.

Spotfire DecisionSite is a visual, analytic application for dynamic, multi-dimensional data analysis. DecisionSite is highly configurable and supports guided analysis in the context of any customer process and data source using analytical tools and built-in data access capabilities.
Use DecisionSite and the R environment to interactively create and test R scripts. These scripts use standard R code and can be parameterized. Using DecisionSite's support for guided analysis, users deploy R scripts as specific end-user data analysis applications. Spotfire DecisionSite users run specific data analysis applications created and deployed to link with an R server. Users are prompted for and enter values for the specific parameters. Results are calculated on the R server and returned to DecisionSite.

The Spotfire Advantage Solution for R provides organizations the ability to deploy algorithms developed in R within a DecisionSite analytic application.

While many statistical tools are freely available, appropriate use of them may be quite difficult for individual users. By linking DecisionSite to an R server through use of the Spotfire Advantage Solution for R, the customer's analytic staff can define specific analysis routines in R and easily make them available to the rest of the organization—for example, various micro-array normalization processes available through Bioconductor.org or a novel high throughput screening
validation routine developed by internal statisticians.