

R and BI – Integrating R with Open Source Business Intelligence Platforms Pentaho and Jaspersoft

David Reinke, Steve Miller

Keywords: business intelligence

Increasingly, R is becoming the tool of choice for statistical analysis, optimization, machine learning and visualization in the business world. This trend will only escalate as more R analysts transition to business from academia. But whereas in academia R is often the central tool for analytics, in business R must coexist with and enhance mainstream business intelligence (BI) technologies. A modern BI portfolio already includes relational databases, data integration (extract, transform, load – ETL), query and reporting, online analytical processing (OLAP), dashboards, and advanced visualization. The opportunity to extend traditional BI with R analytics revolves on the introduction of advanced statistical modeling and visualizations native to R. The challenge is to seamlessly integrate R capabilities within the existing BI space. This presentation will explain and demo an initial approach to integrating R with two comprehensive open source BI (OSBI) platforms – Pentaho and Jaspersoft. Our efforts will be successful if we stimulate additional progress, transparency and innovation by combining the R and BI worlds.

The demonstration will show how we integrated the OSBI platforms with R through use of RServe and its Java API. The BI platforms provide an end user web application which include application security, data provisioning and BI functionality. Our integration will demonstrate a process by which BI components can be created that prompt the user for parameters, acquire data from a relational database and pass into RServer, invoke R commands for processing, and display the resulting R generated statistics and/or graphs within the BI platform. Discussion will include concepts related to creating a reusable java class library of commonly used processes to speed additional development.