

R inside - let R drive your bioinformatic application

Janko Dietzsch, Matthias Zschunke and Kay Nieselt
Junior Research Group Proteomics Algorithms and Simulations
ZBIT (Center for Bioinformatics Tübingen)
University of Tübingen
Sand 14, 72076 Tübingen, Germany
Contact: dietzsch@informatik.uni-tuebingen.de

The emergence of Bioinformatics as a new independent field of science was and is driven by the use of new high throughput methods in life sciences. The huge amount of data produced by these methods causes the need for specialized methods to process, to organize and to visualize the generated data sets. These are the three pillars of the elaborate process of deriving knowledge from the gained information. Based on both its strengths in statistical computing and its rootage in the open source community R has a good chance to play an extraordinary role in Bioinformatics. Projects like Bioconductor already achieved a good acceptance in the research community. But R is a tool primarily designed for professionals and therefore lacks some features that are asked for by users in biological and clinical environments. In particular easy handling via fast reacting GUIs and an excellent visualization is requested by potential users. To deliver improved functionality in these fields we propose to combine R with implementations in an established, well proven language like Java. Java shares with R the independence of the used operating system and has a good support for GUIs, visualization and database connectivity.

We have developed Mayday, short for "MicroarrAY DAta analySis", a plug-in-based platform in Java to analyse and visualize microarray data. It provides Swing-based GUIs and professional visualization features. On the basis of our plug-in architecture we have implemented a connection to R within the Java environment. With this specialized plug-in we offer advanced users the possibility to access R. The chosen way to connect R and our application does not depend on third party libraries and is independent of the underlying operating system. Since it is possible to generalize the implementation details, it can thus be regarded as a case study for integrating R in other software components and can serve as a template for other projects. Last but not least is it a good example of the application of R in a young, interesting and demanding research field - according to the conference motto: "useR!"