

# Using R in a Corporate Environment

John James<sup>1\*</sup>, Francisco Gochez<sup>1</sup>

1. Mango Solutions, UK

\* Contact author: [jjames@mango-solutions.com](mailto:jjames@mango-solutions.com)

**Keywords:** Validation, Assurance, Remoting, Cloud

R is a very powerful data analysis tool. However, its very flexibility often generates the following problem: different users may produce different results on the same data. This is not acceptable in many cases for work in various commercial environments such as the Pharmaceutical and Finance industries –as well as many others - and therefore adoption of R as a mission critical tool is resisted by CIO's in corporate environments.

On the other hand, R is the often the only practicable tool for the end user. This may be because of the advanced nature of the analysis, the size of the data or the short timescale needed for the analysis to be undertaken. At Mango we have worked with numerous large commercial organizations and have implemented a range of solutions to these problems. In this presentation we will discuss some of these approaches and their various pros and cons, and make suggestions about where the R community could usefully contribute to an even wider acceptance and user of R.

In particular, the following points will be covered:

- Validation of code and lockdown. Previous presentations have commented on the ease manipulating an R environment and quality.
- Auditability and reproducibility of results
- Full access to R by end users. Although R is installed in a controlled environment how do users get the access to the command line that is required?
- Supply of the computing resource needed by R: access to large amounts of data, requiring distributed/cloud/grid computing.

The presentation concludes that R is best installed centrally in Corporate Environments with access provided by various bespoke interfaces.

## References

Tony Rossini, David James, "*Open Source Statistical Software in Pharma Development: Case Study with R*", User!2007 (Iowa State University, USA), August 2007  
<http://user2007.org/program/presentations/rossini.pdf>.

Mat Soukup, "*Using R: Perspective of a FDA Statistical Reviewer*", User!2007 (Iowa State University, USA), August 2007  
<http://user2007.org/program/presentations/soukup.pdf>.

Karim Chine, "*Biocep, Towards a Federative, Collaborative, User-Centric, Grid-Enabled and Cloud-Ready Computational Open Platform*," *escience*, pp.321-322, 2008 Fourth IEEE International Conference on eScience, 2008  
<http://biocep-distrib.r-forge.r-project.org/>.