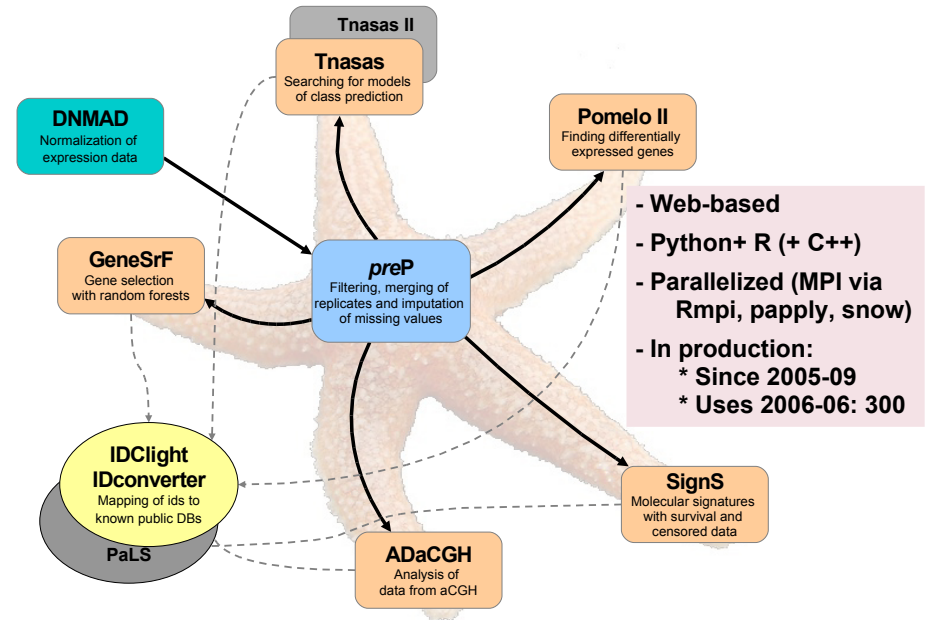# Asterias: an example of using R in a web-based bioinformatics suite of tools

Ramón Díaz-Uriarte, Andrés Cañada, Edward R. Morrissey,
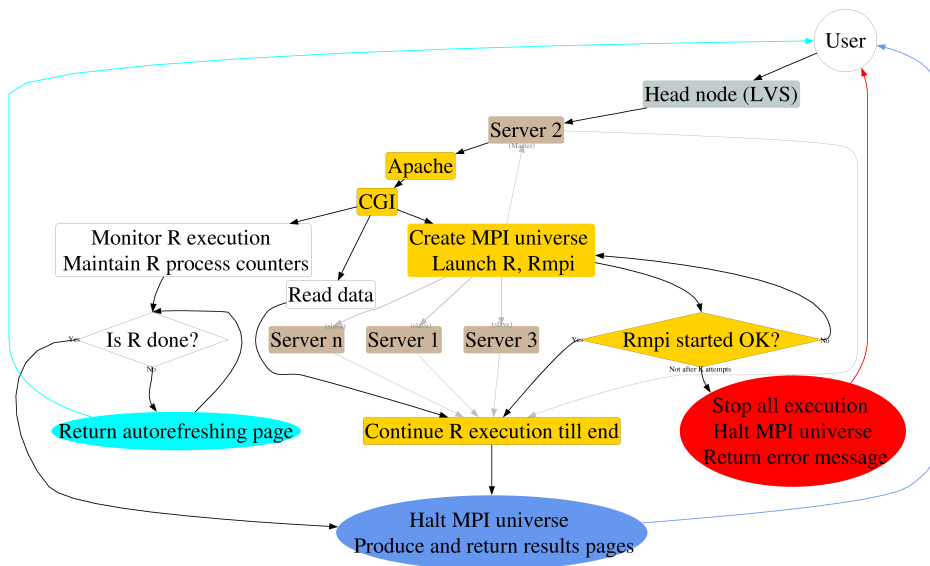Oscar Rueda, Andreu Alibés, David Casado

Statistical Computing Team
Centro Nacional de Investigaciones Oncológicas (CNIO)
(Spanish National Cancer Centre)
rdiaz02@gmail.com
http://ligarto.org/rdiaz

useR! 2006

Asterias tools

- **Web-based**
- **Python+ R (+ C++)**
- **Parallelized (MPI via Rmpi, papply, snow)**
- **In production:**
  - *** Since 2005-09**
  - *** Uses 2006-06: 300**

How it works: overview

Exception and error handling

- CGI and Python pre-processing: catch all user errors
- R: Liberal use of `try`:
  - Orderly exit R (and Rmpi) ($\rightarrow$ whole appl.)
  - Error message (function call), traceback, figure
    - ★ caughtUserError
    - ★ caughtOurError
  - (Few, if any, should ever be user errors: caught before by Python code)

Testing

- Whole system and regression testing: FunkLoad
  http://funkload.nuxeo.org/
  - A suite per application/component
  - Allows simple stress-testing
- Hourly run one FunkLoad test
  - Test head node, servers, MPI
  - Send email if failure
- Other checks (head nodes, NFS servers, each node)

Availability

- Using it: http://www.asterias.info
- Project page: http://bioinformatics.org/asterias
- License: GNU GPL + Affero GPL.

Future

- Simple installation
- Use virtual servers (Linux VServer, Xen) for security and deployment ease
- Further parallelization work: UPC, MPI + OpenMP
- Use of Grid
- Allow deployment as web services