

GIRI Home

GIRI (Generic integration of R into i2b2)

Author: [Bastian Weinlich](#)

Short Description

The project follows a similar goal as the "R Engine Cell" of the University of Pavia and [rgate \(HERON\)](#) of the University of Kansas: To make arbitrary R functions available within i2b2. In contrast to these approaches, it is very simple to add new statistical functionality with GIRI. To do so, it is sufficient to write an R script and an optional XML configuration file to add what is known as a "Scriptlet" (a kind of an addon) that comprises one or more R functions. In particular, it is not necessary to implement an i2b2 webclient plugin for every newly added scriptlet individually. The end user simply selects the desired scriptlet from a drop down list from a generic i2b2 webclient plugin and has then several options to customize the input data (e.g. drag and drop patient sets / concepts). Afterwards, he clicks on "View Results" to start computations. The following results page can - depending on the scriptlet - consist of numerical / textual values, tables and plots. Additionally, all data can be exported as files.

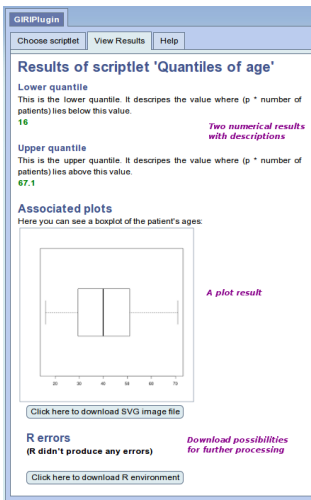
Examples

A simple scriptlet could compute two quantiles of age from a previously selected patient set. The end user can define the p values. Here is a step-by-step guide to using the scriptlet:

Specify the input data:

The screenshot shows the GIRIPlugin web interface. At the top, there are tabs for "Choose scriptlet", "View Results", and "Help". Below the tabs, a dropdown menu is open, showing "Quantiles of age" selected. To the right of the dropdown, there is a note: "Choose between several previously written scriptlets". Below the dropdown, the title "Quantiles of age" is displayed. Underneath, there is a prompt: "Please select a patient set and fill in two values for p." The "Drag and drop fields:" section contains a red button labeled "Click here to clean fields" and a note: "Define input data by dragging and dropping patient sets and concepts". There are two draggable boxes: "Circula-Male@18:51:46 [4]" and "Concept 1". Below these, there is a "Patient Set 2" box. The "Additional input settings" section has a note: "Type in additional parameters to customize the computation". It includes two input fields: "p for lower quantile" with a value of "0.02" and "p for upper quantile" with a value of "0.98".

Afterwards click on "View Results" to start the computation:



The scriptlet consists of two files that have to be stored on the server where the GIRI Cell is running. The first one is the R script including the following R code:

mainscript.r

```
GIRI.output["Lower quantile"] <- quantile(GIRI.patients[[1]]$age_in_years_num, as.numeric(GIRI.input[["p for lower quantile"]]))
GIRI.output["Upper quantile"] <- quantile(GIRI.patients[[1]]$age_in_years_num, as.numeric(GIRI.input[["p for upper quantile"]]))
boxplot(GIRI.patients[[1]]$age_in_years_num, horizontal=T)
```

Accessing input and output is possible by using special naming conventions (GIRI.patients, GIRI.input, GIRI.output...). Plots are always displayed.

The second file configures the R scriptlet through XML code:

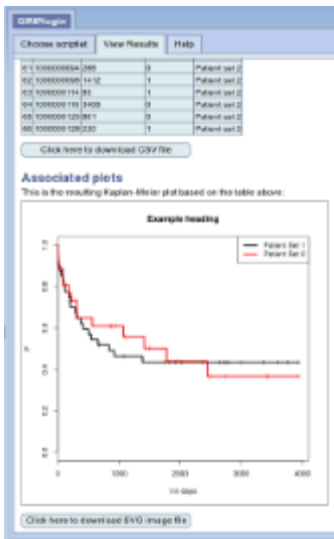
config.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<giriconf:Rscriptlet xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:giriconf="http://www.i2b2.org/xsd/cell/giriconf/1.0/">
  <settings>
    <title>Quantiles of age</title>
    <description>Please select a patient set and fill in two values for p.</description>
    <passROutput>false</passROutput>
    <passRErrors>true</passRErrors>
    <plotDescription>Here you can see a boxplot of the patient's ages:</plotDescription>
  </settings>
  <additionalInputs>
    <input>
      <name>p for lower quantile</name>
      <description>Fill in a numeric value between 0.0 and 1.0, e.g. 0.05</description>
      <type>text</type>
      <lines>1</lines>
    </input>
    <input>
      <name>p for upper quantile</name>
      <description>Fill in a numeric value between 0.0 and 1.0, e.g. 0.95</description>
      <type>text</type>
      <lines>1</lines>
    </input>
  </additionalInputs>
  <customOutputs>
    <output>
      <name>Lower quantile</name>
      <description>This is the lower quantile. It describes the value where (p * number of
patients) lies below this value.</description>
    </output>
    <output>
      <name>Upper quantile</name>
      <description>This is the upper quantile. It describes the value where (p * number of
patients) lies above this value.</description>
    </output>
  </customOutputs>
</giriconf:Rscriptlet>
```

Settings, input variables and outputs are defined in this configuration file. For more information about creating scriptlets see the [notes for scriptlet developers](#).

The following slide show depicts a scriptlet that computes a Kaplan Meier plot:

Kaplan Meier Example



Kaplan-Meier-input-table
Following data was computed from database (0 means censored, 1 means not censored):

Patient	Time in days	Censored	Group
1	0	1	Patient set 1
2	100	0	Patient set 1
3	100	1	Patient set 1
4	150	0	Patient set 1
5	150	1	Patient set 1
6	200	1	Patient set 1
7	200	0	Patient set 1
8	250	0	Patient set 1
9	250	0	Patient set 1
10	250	1	Patient set 1
11	300	1	Patient set 1
12	300	1	Patient set 1
13	300	0	Patient set 1
14	350	1	Patient set 1
15	350	1	Patient set 1
16	350	0	Patient set 1
17	350	0	Patient set 1
18	350	0	Patient set 1
19	350	0	Patient set 1
20	350	0	Patient set 1
21	350	1	Patient set 1
22	350	1	Patient set 2

Please choose a GRI scriptlet:
Kaplan-Meier

Kaplan-Meier
Usage information: Drag and drop exactly one risk concept (e.g. a cancer diagnosed on the "Concept 1" field. Afterwards a "Concept 2" field will appear. Please use this "Concept 2" field to drag and drop exactly one target concept (e.g. deceased) on it. Then drag and drop as many patient sets you wish to compare among each other on the left fields. In doing so make sure that all members of this set have the risk concept. Example: Use the query tool to obtain patient set A "50 female patients having breast cancer and being older than 50 years" and patient set B "All female patients having breast cancer and being younger than 50 years". Drag and drop these patient sets from the "Previous Query" box below to the left fields below in the scriptlet window. Then drag and drop "Breast cancer" on "Concept 1" as risk concept and "Deceased" on "Concept 2" as target concept. "Concept 3" and further concept fields will be ignored here. Finally specify the plot parameters below and click on "View result".

Drag and drop fields:

- Circle 1 Male@19:51:48:14
- Circle 2 Female@19:51:22
- Patient Set 3
- Circulatory system
- Digestive system
- Concept 3

Additional input settings
Plot heading: Example heading
Colorful or black-white plot graphics? Colorful

Please choose a GRI scriptlet:

- Overview of ages
- Example 1: Field work
- Kaplan-Meier Modified
- Frequency of child concept observations
- Kaplan-Meier**
- Example 2: Testing R output
- Example 4: Additional input and custom output
- Example 3: Access and view data
- Laboratory value distributions
- Example: Scriptlet: All features

Downloads

The download package contains the GRI Cell, the GRI webclient plugin, install notes, notes for scriptlet developers and example scriptlets. This version is dedicated for i2b2 version 1.7.02 (does not work with 1.7.01!).

[giri_i2b2_project_v1.1.zip](#)