DESIRR

Semantic platform for organizing, sharing and reusing R functions

Pascal Neveu\textsuperscript{a}, Yuan Lin\textsuperscript{a}, Caroline Domerg\textsuperscript{a}, Juliette Fabre\textsuperscript{a}, Alexandre Granier\textsuperscript{a}, Anne Tireau\textsuperscript{a}, Vincent Negre\textsuperscript{a}, Isabelle Mirbel\textsuperscript{b}, Olivier Corby\textsuperscript{c}, Catherine Faron-Zucker\textsuperscript{b}

\texttt{pascal.neveu@supagro.inra.fr}

\textsuperscript{a} INRA, Montpellier, France
\textsuperscript{b} I3S, University of Nice, Sophia Antipolis, France
\textsuperscript{c} INRIA, Sophia Antipolis Méditerrane, Sophia Antipolis, France
Overview

- R functions
  - Great success in different scientific domains
  - High production of R functions

- Authors from different sites, with authors turn over

- Challenge: Open “data” (*function*)
  - Share and reuse
Context

- Objectives
  - Store and organize available R functions
  - Give an easy and long-term access

- How
  - Create **formal description** for R functions based on **ontological approach**
    - Define concepts and relations between concepts *(Domain expertise)*
    - Controlled vocabulary for management and search facility
    - Build a function repository with powerful reasoning and search capabilities

- Technologies
  - Methods and tools of **Semantic Web** *(W3C)* *(Knowledge organization and management)*
Leaf_Growth_Analysis <- function(para1, para2 ...) {
    ...
}

hasAuthor
Ontology (Definition)

```
Leaf_Growth_Analysis <-
function(para1, para2 ...)
{
    ...
}
```

**Triplet representation**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;#Leaf_Growth_Analysis.R&gt;</td>
<td>&lt;#hasAuthor&gt;</td>
<td>&lt;#Vera&gt;</td>
</tr>
</tbody>
</table>

hasAuthor
Ontology (Definition)

Leaf_Growth_Analysis <- function(para1, para2 ...) {
    ...
}

Triplet representation

Subject

Predicate

Object

hasAuthor

Graph representation

Leaf_Growth_Analysis.R

hasAuthor

Vera

Leaf_Growth_Analysis.R

hasAuthor

Vera
Ontology (Definition)

- **dataframe**
  - **hasArgument**: `dataframe`
  - **hasArgument**: `xxxxx`
  - **hasEmail**: vera@inra.fr
  - **hasAuthor**: #RFunction
  - **hasAuthor**: #Leaf_Growth_Analysis.R
  - **hasAuthor**: #Person
  - **hasAuthor**: #Vera

- **hasEmail**: vera@inra.fr

```
<#RFunction><#type><#Person><#hasEmail>vera@inra.fr
<#Leaf_Growth_Analysis.R> <#hasAuthor> #Vera
<#para1> <#hasArgumemt> dataframe
<#para1> <#dataType> x
<#para1> <#defaultValue> x
```
Ontology (Definition)
Ontology (Model-Data)
Ontology (Model-Data)
R function ontology (overview)

- General desc, Detailed info, Relationship between R functions

![Diagram of R function ontology]

- #name
  - Has name
- #description
  - Has description
- #date
  - Has creation date
  - Has repository date
- #documentation
  - Has documentation
- #Rpackage
  - Requires R package
- #intention
  - Has Intention
- #type
  - Has type
  - Has default value
- Value
  - Has value
  - Calls R core function
- Person
  - Has author
  - Has contact
  - Has audience
  - Belongs to
  - Group
- Argument
  - Has argument
  - Has type
  - Has description
- #defaultValue
  - Has default value
- #RcoreFunction
  - Calls local R function
  - Could be used after
  - Could be used before
  - Is adapted from
  - Is a new version of
  - Looks like
General description
name, creationDate, desc, author, contact, audience, documentation
R function ontology

- Detailed information

  requiresPackage, hasArgument, hasValue, hasIntention, etc.

Graph:

- R Function
  - Requires R package
    - #Rpackage
  - Has Intention
    - #intention
  - Has argument
    - Has type
      - Has description
    - Has value
      - #type
        - #description
        - #defaultValue
  - Has default value
    - Has type
      - Has description
      - Has default value
R function ontology

- **Relationship between R functions**
  - callsRCoreFunc, callsLocalFunc, couldBeUsedBefore/After, looksLike, isAdaptedFrom, isNewVersionOf.

![Diagram of R function relationships](image-url)
R function ontology

- **Relationship between R functions**
  - callsRCoreFunc, callsLocalFunc, couldBeUsedBefore/After, looksLike, isAdaptedFrom, isNewVersionOf.

Creation of the call graph
Application (architecture)

- **Application web** (works with a web browser)
- **R terminal** (parser and validate the uploaded R files)
- **Description storage** in the **server** and in a specific **data base**
- **Version** and **concurrency control**

Diagram:
- Web browser
- Web application for R Annotations (PHP/Apache Server)
- RDF Annotations
- R Functions
- Datasets, doc...
- Versioning (Subversion)
- Triple store (CORESE INRIA)
Application (use case)

- New description creation
Application  (use case)

- New description creation

Upload a R func file  
Parse & validate by R client  
Fill the desc form
Application (use case)

- New description creation

Upload a R func file  Parse & validate by R client  Fill the desc form

1 - General

Fields followed by * should be filled!

Name of the function *: add.R

The name should be of the following form: MyFunction.R

Description *:

Multiple selection or unselection: use <Ctrl>

Author(s):
- Anne.Pellegrino
- Anne.Tireau
- Benoit.Boussuge
- Bertrand.Muller
- Caroline.Domerg
- Christian.Fournier

Person(s) to contact:
- Christian.Fournier
- Christine.Granier
- Eric.Lebon
- Bertrand.Muller
- Vincent.Negre
- Anne.Pellegrino
Application (use case)

- New description creation

Upload a R func file  Parse & validate by R client  Fill the desc form

Argument description

Describe all the different arguments of the function:

Argument 1
Name: [ ] Type: [ ] Description: [ ]

Add argument

Value description

Describe all the different values of the function:

Value 1
Name: [ ] Type: array Description: [ ]

Add value
Application (use case)

- New description creation

Upload a R func file
Parse & validate by R client
Fill the desc form

4 - Relations with other functions of the repository

<table>
<thead>
<tr>
<th>Is a new version Of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Add</td>
</tr>
<tr>
<td>Show List</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Is adapted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>Could be used after</td>
</tr>
<tr>
<td>Could be used before</td>
</tr>
<tr>
<td>Calls</td>
</tr>
<tr>
<td>Looks like</td>
</tr>
</tbody>
</table>
Application (use case)

- New description creation

Upload a R func file
Parse & validate by R client
Fill the desc form
Save the R func & its desc

---

4 - Relations with other functions of the repository

Is a new version Of

<table>
<thead>
<tr>
<th>Is adapted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>Could be used after</td>
</tr>
<tr>
<td>Could be used before</td>
</tr>
<tr>
<td>Calls</td>
</tr>
<tr>
<td>Looks like</td>
</tr>
</tbody>
</table>
Application (use case)

- Function consultation
Application (use case)

- Function consultation

**Example:**
isDedicatedTo 'Phenody' and hasIntention of 'Visualization'
**Application (use case)**

- Function consultation

**Input search conditions**

**Generalize and run SPARQL query**

**Example:**

isDedicatedTo 'Phenody' and hasIntention of 'Visualization'

**PREFIX OntologyR: select ?fonction ?description where { ?fonction OntologyR:isDedicatedTo OntologyR:Phenody ?fonction OntologyR:hasIntention OntologyR:Visualisation ?fonction OntologyR:hasDescription ?description}**
**Application** (use case)

- **Function consultation**

  - **Input search conditions**
  - **Generalize and run SPARQL query**
  - **Show result list**

---

**There are 12 functions matching your request:**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRTempPlot_Caroline.Domerg.R</td>
<td>This function generates a pdf file with the response curves of the growth speed to the temperature, one plot per genotype and the coefficients and line on the regression are displayed. A csv file with the parameters of the regressions is also returned.</td>
</tr>
<tr>
<td>CRpsiPlot_Caroline.Domerg.R</td>
<td>This function generates a pdf file with the response curves of the growth speed to the PSI, one plot per genotype and the coefficients and line on the regression are displayed. A csv file with the parameters of the regressions is also returned.</td>
</tr>
<tr>
<td>ancCRpsi_Caroline.Domerg.R</td>
<td>This function performs and displays covariance analysis on the response curves of the growth to the psi. It compares the response of a list of genotypes to a probe genotype to study the effect of the genotype on the response of the growth.</td>
</tr>
<tr>
<td>CRvpdPlot_Caroline.Domerg.R</td>
<td>This function generates a pdf file with the response curves of the growth speed to the leaf-EPD, one plot per genotype and the coefficients and line on the regression are displayed. A csv file with the parameters of the regressions is also returned.</td>
</tr>
<tr>
<td>LERvalidation_Vera.Georgescu.R</td>
<td>The general function of visualisation, automatic and manual correction of the Leaf Elongation Rate kinetics measured on the Phenodyn platform. This function runs on R version 2.6.2.</td>
</tr>
</tbody>
</table>
Application (use case)

- Function consultation

- Input search conditions
- Generalize and run SPARQL query
- Show result list
- Consult & modify descriptions

...
Application (use case)

- Function consultation
  - Input search conditions
  - Generalize and run SPARQL query
  - Show result list
  - Consult & modify descriptions

Call graph of the current function

Relations with other functions
Application (inference)

Leaf_Growth_Computation.R

Could be used after

Leaf_Growth_Analysis.R

Calls local R function

Growth_Curve_Fitting.R

Calls local R function

Leaf_Growth_Modeling.R
Application (inference)

Leaf_Growth_Computation.R

Could be used after

Leaf_Growth_Analysis.R

Could be used before

INVERSE OF

Growth_Curve_Fitting.R

Calls local R function

Leaf_Growth_Modeling.R

Calls local R function
Could be used after

Leaf_Growth_Computation.R

Could be used before

INVERSE OF

Calls local R function

TRANSITIVITY

Calls local R function

Leaf_Growth_Analysis.R

Calls local R function

Growth_Curve_Fitting.R

Leaf_Growth_Modeling.R
Application (use case)

- Inference example 1
  - Upload Add.R
  - Upload Add2.R
    Add2.R adaptedFrom Add.R
  - Upload Add3.R
    Add3.R adaptedFrom Add2.R
Application (use case)

- Inference example 1
  - Upload Add.R
  - Upload Add2.R
    Add2.R adaptedFrom Add.R
  - Upload Add3.R
    Add3.R adaptedFrom Add2.R

Add3.R is adapted from which function(s)?
Application (use case)

- Inference example 1
  - Upload Add.R
  - Upload Add2.R
    *Add2.R adaptedFrom Add.R*
  - Upload Add3.R
    *Add3.R adaptedFrom Add2.R*

Add3.R is adapted from which function(s) ?

*Is adapted from: add2_Yuan.Lin.R*
**Application (use case)**

- **Inference example 1**
  - Upload Add.R
  - Upload Add2.R
    **Add2.R adaptedFrom Add.R**
  - Upload Add3.R
    **Add3.R adaptedFrom Add2.R**

Add3.R is adapted from which function(s)?

*Sequence*

Is adapted from: add2_Yuan.Lin.R
Application (use case)

- Inference example 1
  - Upload Add.R
  - Upload Add2.R
    Add2.R \textit{adaptedFrom} Add.R
  - Upload Add3.R
    Add3.R \textit{adaptedFrom} Add2.R

Add3.R is adapted from which function(s) ?
Application  (use case)

- Inference example 2
  - Upload Func1
  - Upload Func2
    Func2  couldBeUsedBefore  Func1
  - Upload Func3
    Func3  couldBeUsedAfter  Func2
Application (use case)

- Inference example 2
  - Upload Func1
  - Upload Func2
    Func2 couldBeUsedBefore Func1
  - Upload Func3
    Func3 couldBeUsedAfter Func2

What are functions that could be used after Func2
Inference example 2
- Upload Func1
- Upload Func2
  \textbf{Func2 couldBeUsedBefore} Func1
- Upload Func3
  \textbf{Func3 couldBeUsedAfter} Func2

What are functions that could be used after Func2

\textbf{Func3}
Application (use case)

- Inference example 2
  - Upload Func1
  - Upload Func2
  - Upload Func3

What are functions that could be used after Func2

- Func1
- Func2
- Func3
Application  (use case)

- Inference example 2
  - Upload Func1
  - Upload Func2  Func2  couldBeUsedBefore  Func1
  - Upload Func3  Func3  couldBeUsedAfter  Func2

What are functions that could be used after Func2

- Func3
- Func1
Conclusion

- Users find this repository relevant (efficient search, easy annotating)

- Semantic Web tools allow reasoning for an “intelligent” repository

- Models and software are easy to adapt:
  - for other research fields
  - for other programming languages
  - for mathematical models?
THE END

Thank you

pascal.neveu@supagro.inra.fr
yuan.lin@supagro.inra.fr