

# Using ontologies for function management

Caroline Domerg, Juliette Fabre and Pascal Neveu

O. Corby  
C.Faron-Zucker  
E.Gennari  
A. Granier  
I. Mirbel  
V. Negre  
A. Tireau

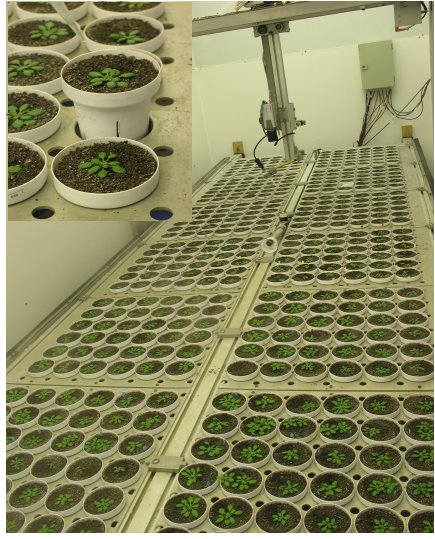
**Context**

**Semantic Web tools**

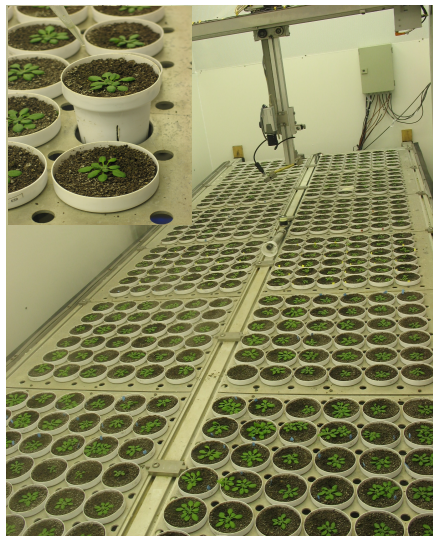
**Ontology description**

**Web interface overview**

**Conclusion**



- Plant adaptation to climatic change
- Environmental stresses X several species
- High-throughput phenotyping
- Databases

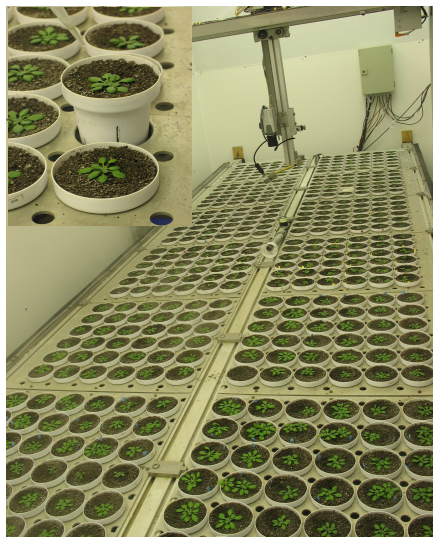


- Plant adaptation to climatic change
- Environmental stresses X several species
- High-throughput phenotyping
- Databases



Experiment  
monitoring

Data analysis  
and modeling



- Plant adaptation to climatic change
- Environmental stresses X several species
- High-throughput phenotyping
- Databases

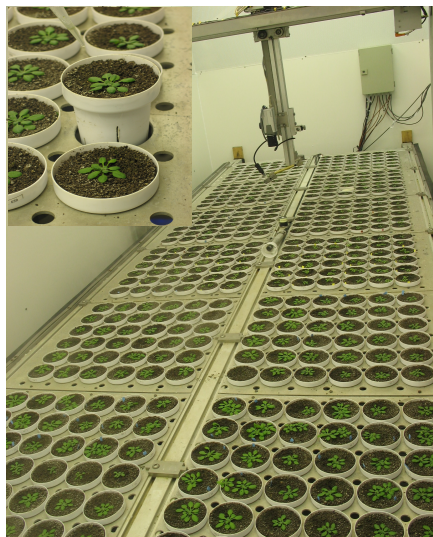


Experiment  
monitoring

Data analysis  
and modeling



**High production  
of R functions**



- Plant adaptation to climatic change
- Environmental stresses X several species
- High-throughput phenotyping
- Databases



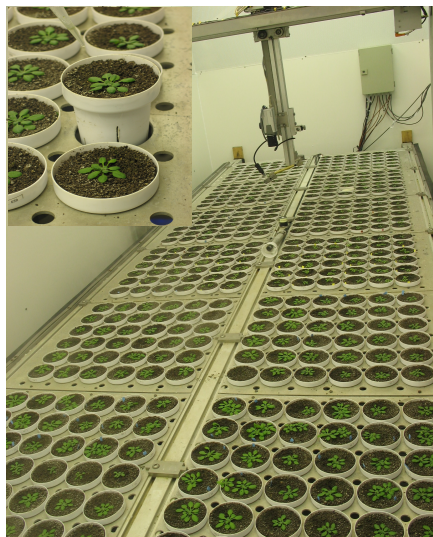
Experiment  
monitoring

Data analysis  
and modeling



Many authors  
and turnover

**High production  
of R functions**



- Plant adaptation to climatic change
- Environmental stresses X several species
- High-throughput phenotyping
- Databases



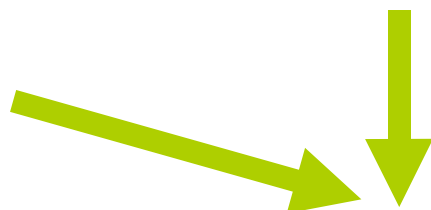
Experiment  
monitoring

Data analysis  
and modeling



Many authors  
and turnover

**High production  
of R functions**



**How to share, capitalize, organize and valorize  
these functions?**

## AIMS

Store and organize the functions  
Give an easy and long-term access

## IDEAS

Create an ontology to describe R functions  
Provide a new kind of repository with reasoning and powerful search tools

## TOOLS

W3C Semantic Web technologies



# Ontology

Formal description of concepts and relations between concepts

*Examples of concept:* **Rfunction** **Argument** **Person**

*Examples of relation:* **hasArgument** **isANewVersionOf**

→ Provides a **controlled** vocabulary

→ Designed to be understood by **computers**

RDF, RDFS and OWL: standard tools to write ontologies

# RDF

**Resource:** documents, images, programs, etc

**Description:** attributes, properties and relations

**Framework:** model, language and syntaxes (XML) for these descriptions

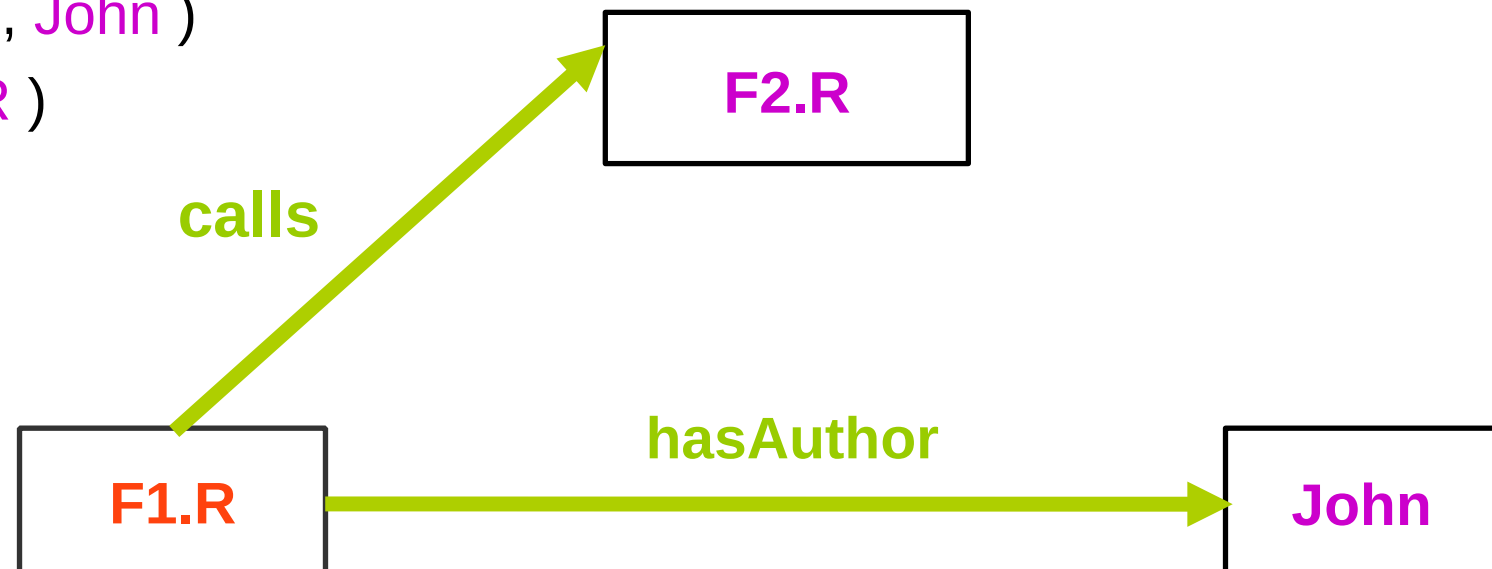
# RDF

An example about R function descriptions:

- R function attributes and properties
- Relations between R functions

( F1.R , hasAuthor , John )

( F1.R , calls , F2.R )



RDF is a semantic graph model

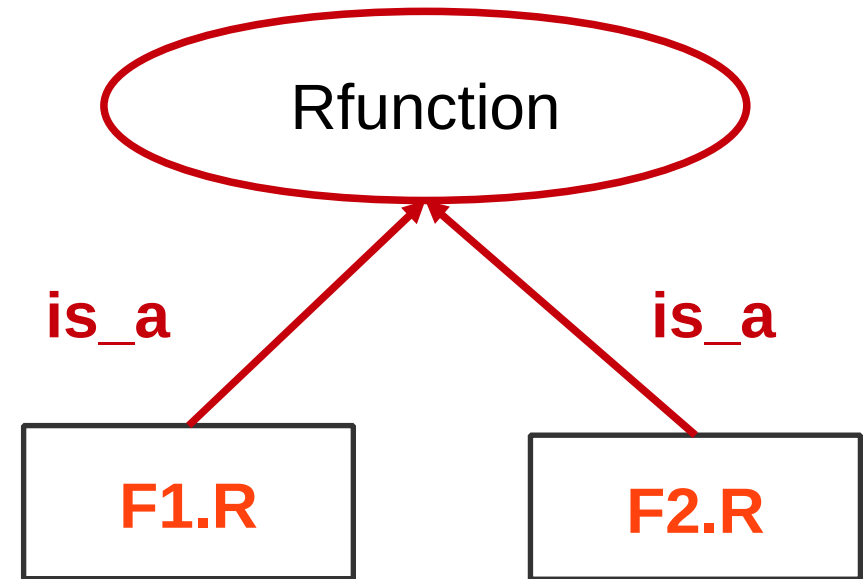
# RDFS

## RDF Schema

Provides elements to **structure** RDF resources such as:

- Class hierarchy
- Property restrictions (domain, range)

*Example of class: the Rfunction class*



➔ F1.R and F2.R inherit the properties and attributes of the Rfunction class

# OWL Ontology Web Language

- Built on top of RDF
- Allows to define rules: transitivity, symmetry, inverse of, etc
- ➔ Provides powerful description of concepts and their relationships

*Example of OWL rule: INVERSE OF*



# Ontology querying

## RDF/OWL files

### Ontology and annotations

```
<owl:ObjectProperty rdf:about="#couldBeUsedAfter">  
  <rdfs:range rdf:resource="#Rfunction"/>  
  <rdfs:domain rdf:resource="#Rfunction"/>  
  <owl:inverseOf rdf:resource="#couldBeUsedBefore"/>  
</owl:ObjectProperty>
```

## SPARQL

Query

Language

for RDF/OWL

# Ontology querying

## RDF/OWL files

### Ontology and annotations

```
<owl:ObjectProperty rdf:about="#couldBeUsedAfter">  
  <rdfs:range rdf:resource="#Rfunction"/>  
  <rdfs:domain rdf:resource="#Rfunction"/>  
  <owl:inverseOf rdf:resource="#couldBeUsedBefore"/>  
</owl:ObjectProperty>
```

## SPARQL

Query

Language

for RDF/OWL

## CORESE (*Inria – Edelweiss*)

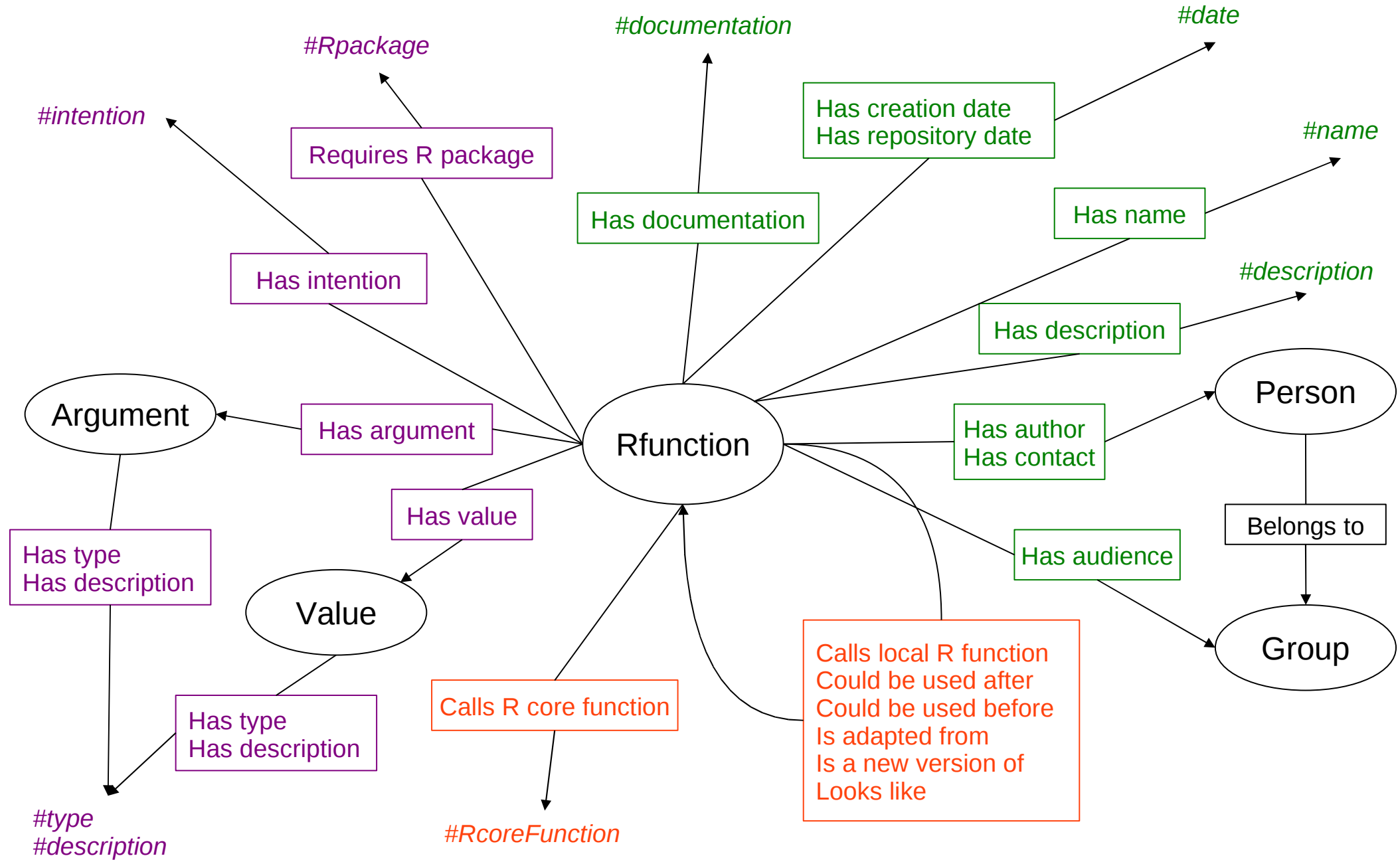
Software and engine to run:

- rules (transitivity, etc) and to infer
- SPARQL queries



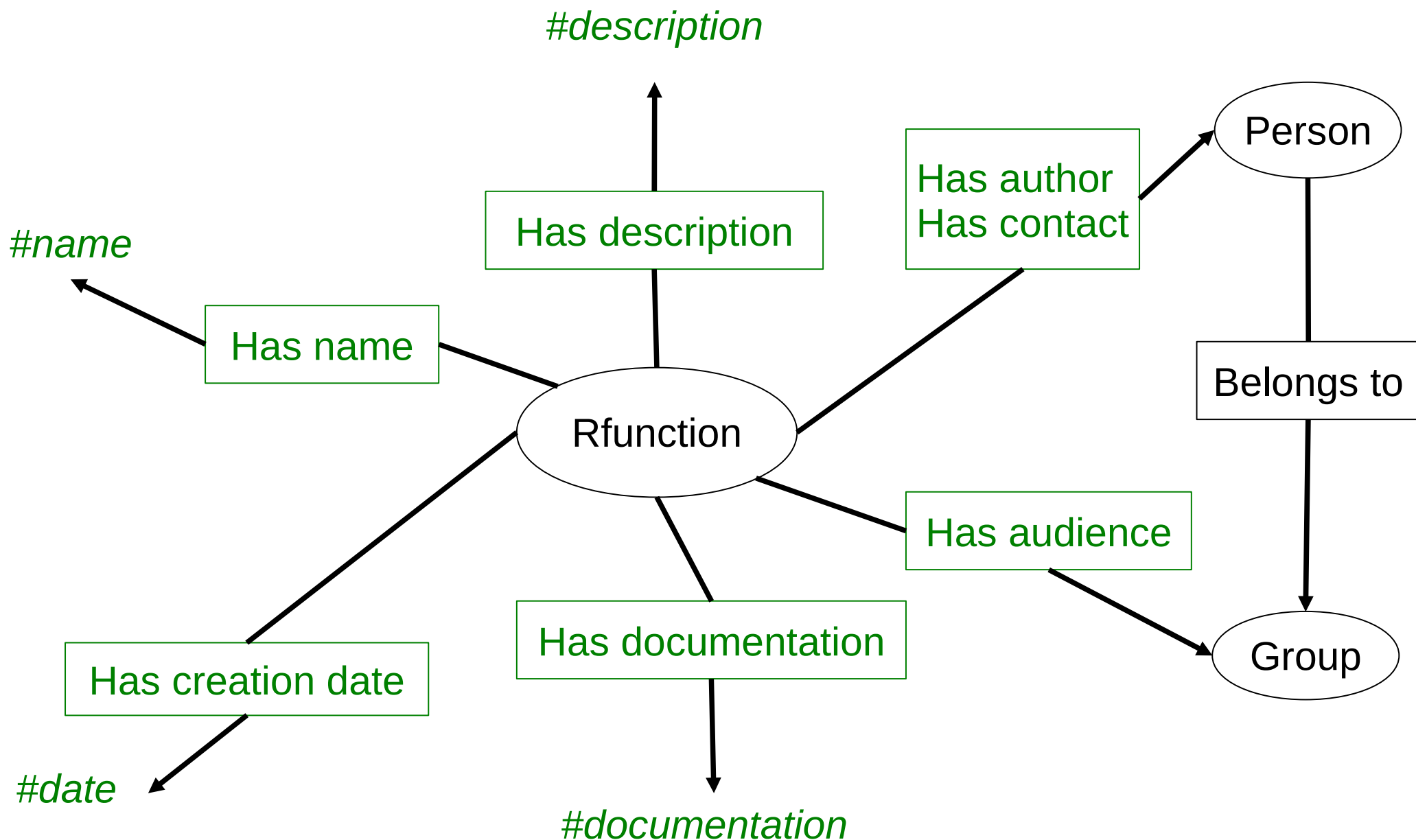
**Query  
results**

# Global view

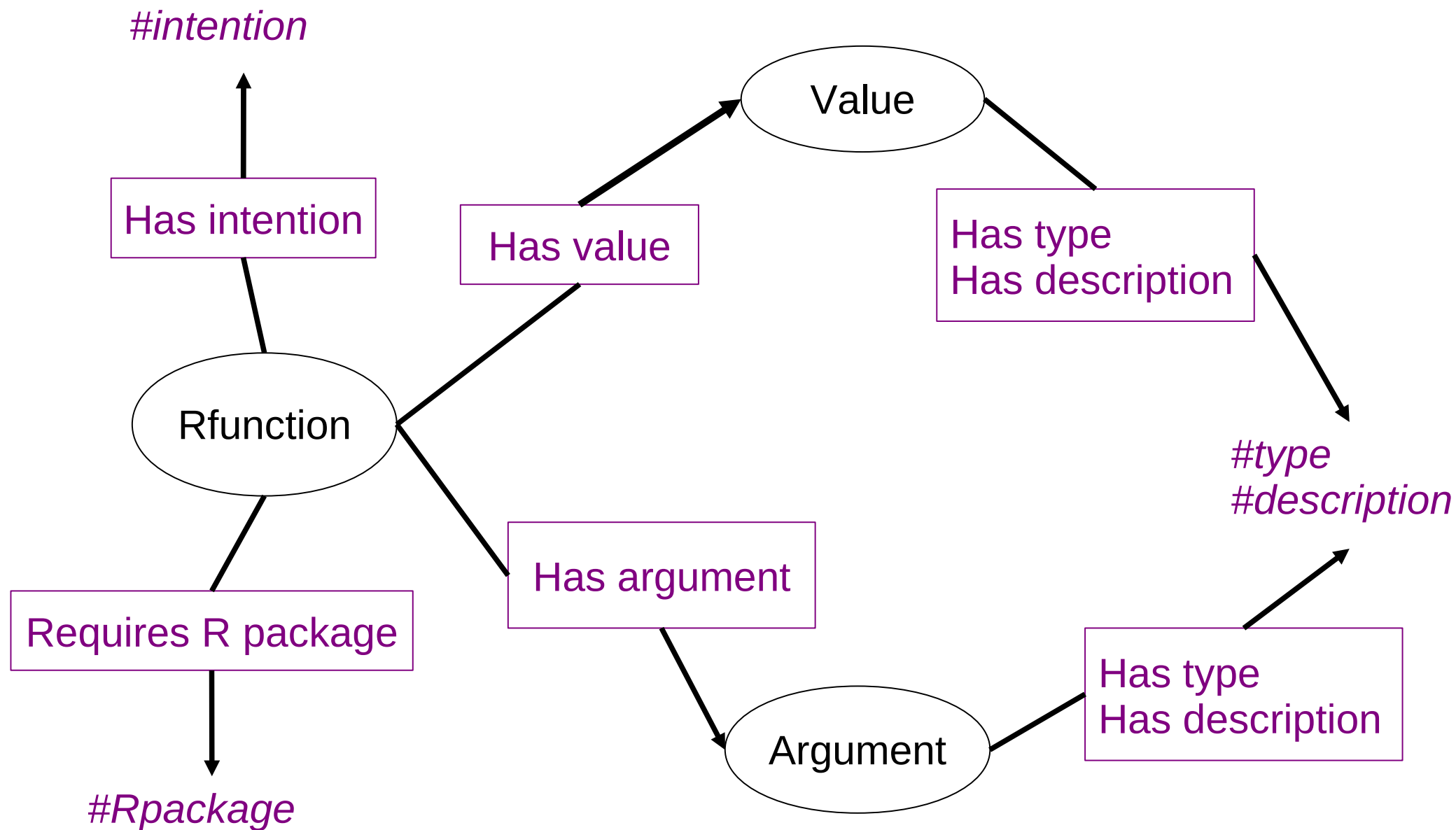




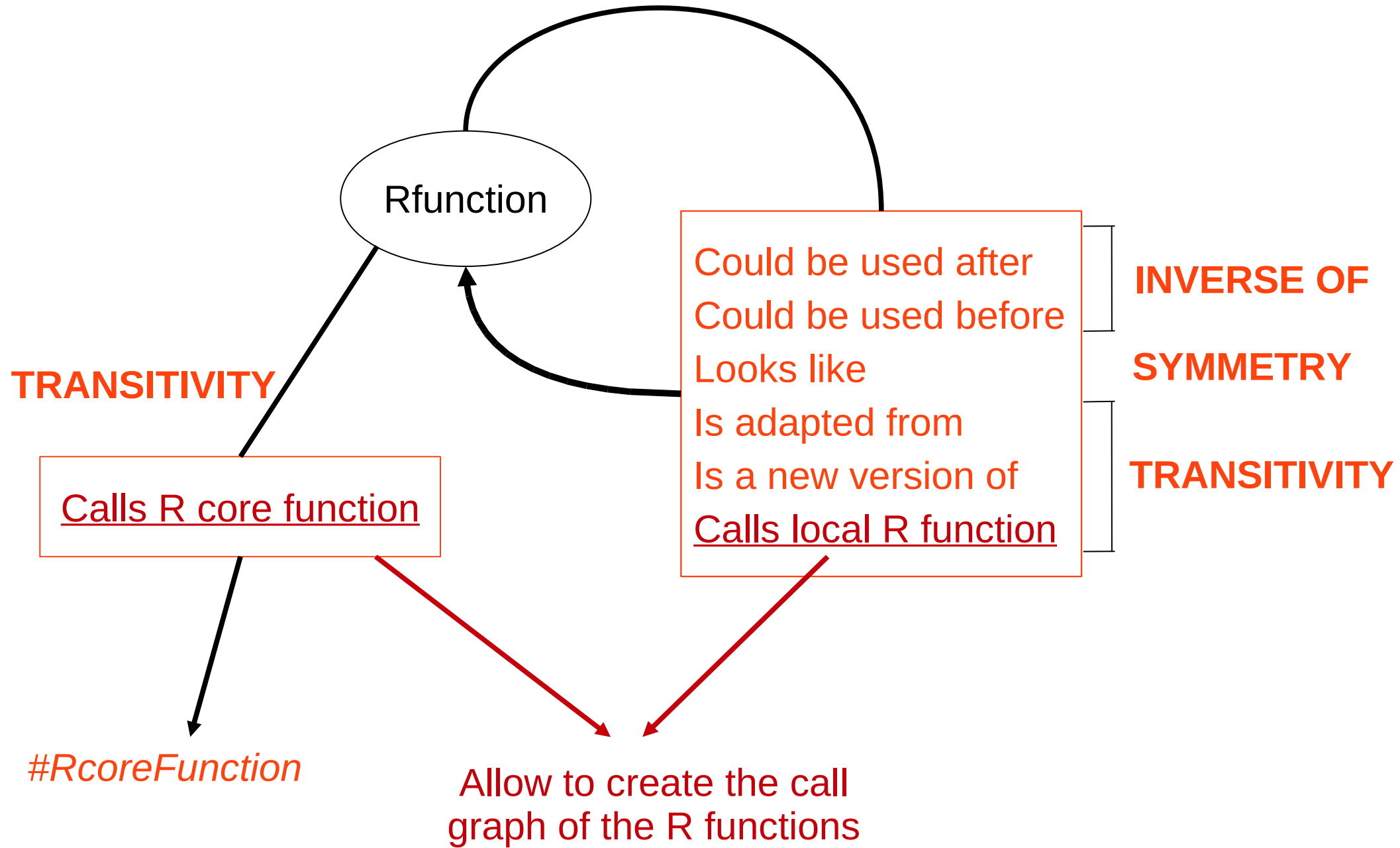
# General description of R functions



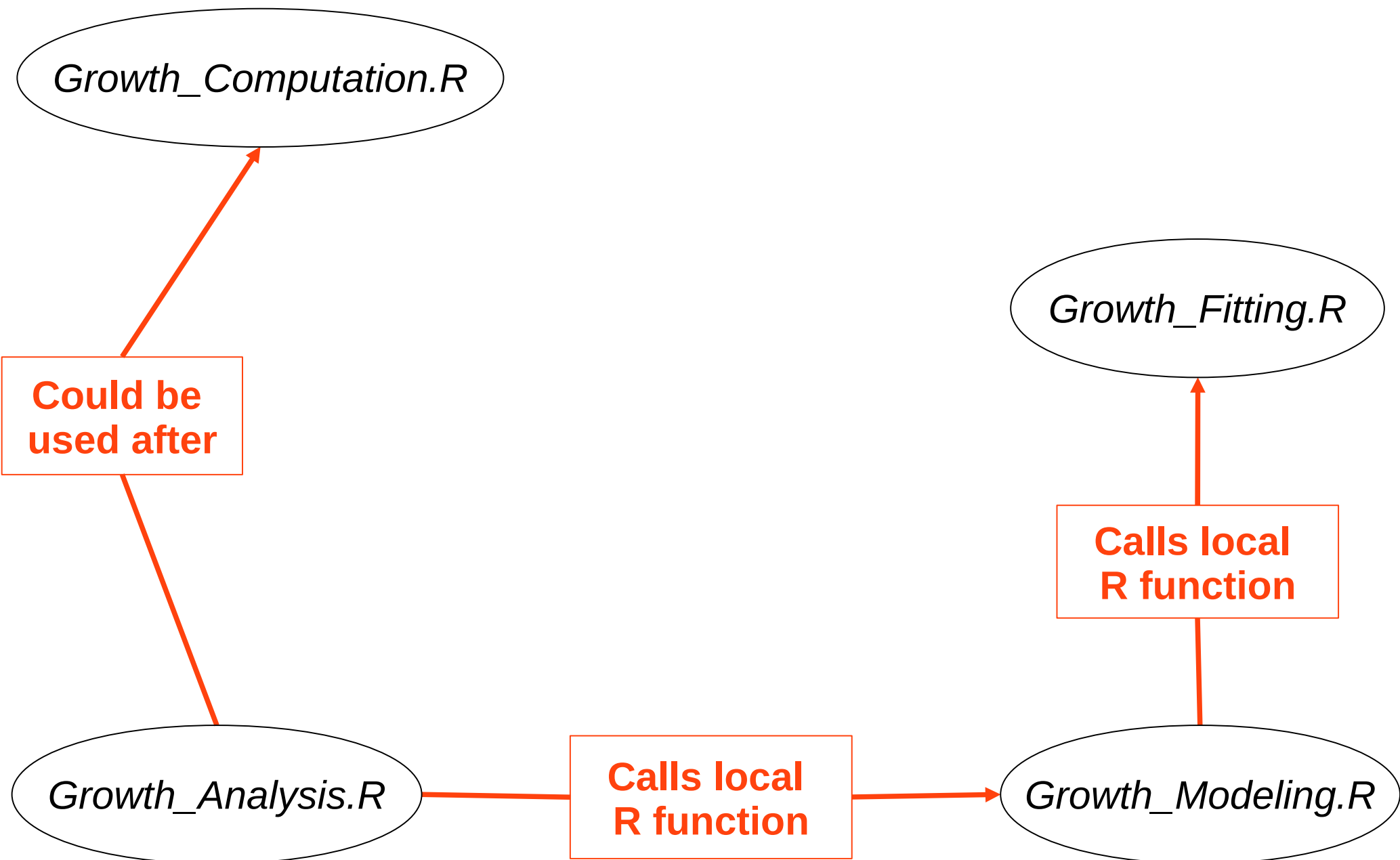
# Detailed description of R functions



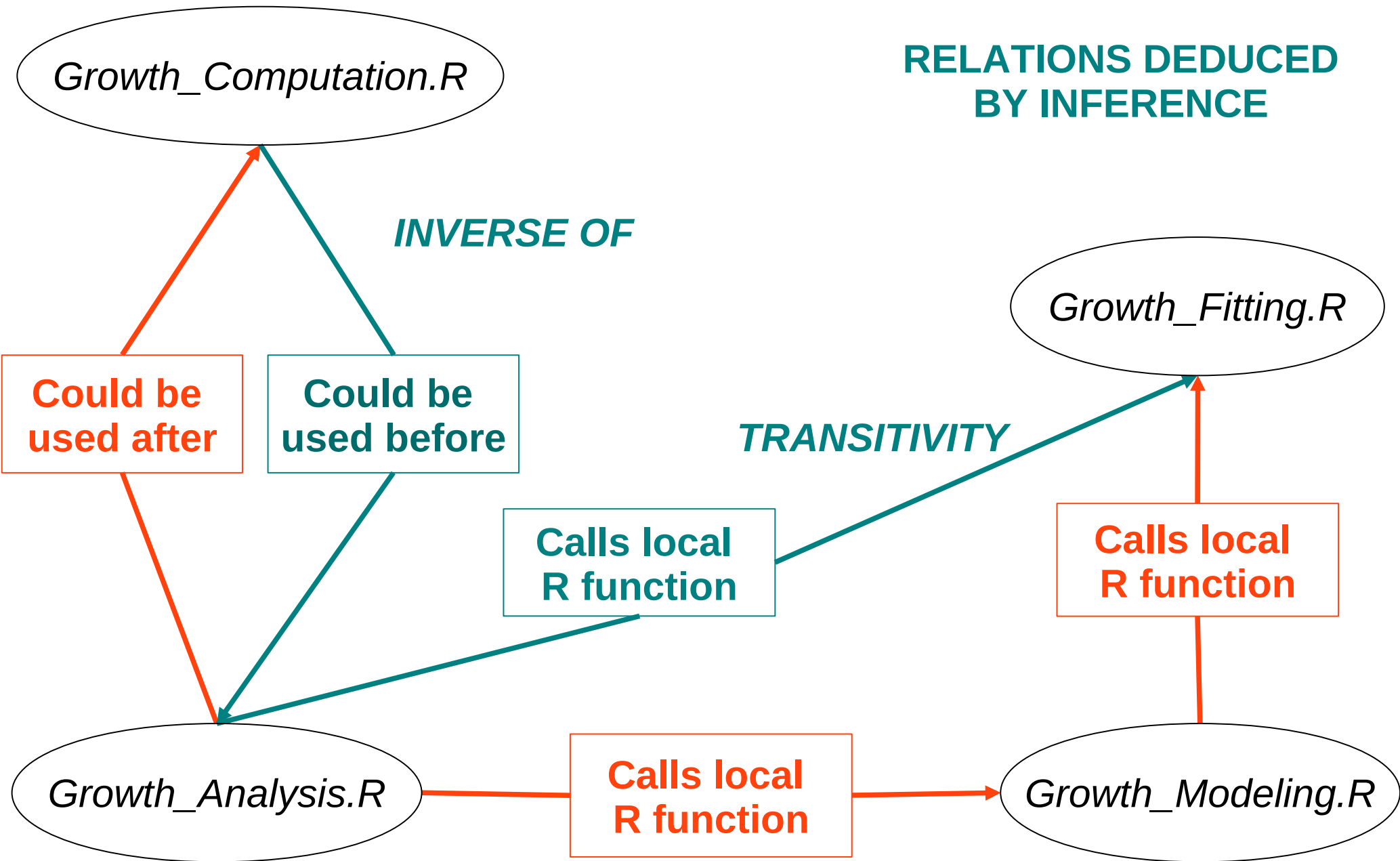
# Relations with other R functions



## Relations with other R functions




# Relations with other R functions



# GUI for edition and creation of annotations

- A few minutes thanks to pre-filled forms
- Generation and storage of OWL file


## 1 - General

 Fields followed by \* should be filled!

Name of the function \* :

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

Description \* :

 Multiple selection or unselection: use <Ctrl>

Author(s):


Person(s) to contact :

Creation date (ex: 2010-11-26) \* :

## 2 - Uploads

Upload the R script (.R):

## 3 - Detailed description

 All the following informations are optional:

Audience:

Information system(s) concerned:


Intention(s) of the function:

Execution time of the function (short, medium, long):

R package(s) required separated by ',' (eg: ade4; lattice):

Main R functions called in the function separated by ',' (eg: lm; curve; nls):

### Argument description

 Describe all the different arguments of the function:

Argument 1 Name:  Type:  Description:

# Function consultation card

## Vera.Georgescu\_LERvalidation.R

 [Download R files](#)

### Description

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

### Arguments

#### ■ LERvalidation.R\_graph

Type: logical

Description: boolean for graphic mode (for manual correction)

#### ■ LERvalidation.R\_finnut

Type: scalar

Description: the hour of end of the night

### Audience

Ecophysiologicalist

### Authors

Vera.Georgescu

### Contacts

Vincent.Negre

### Creation date

2008-02-01

→ Information about the function (author, arguments, intentions, etc)

→ Download of R function and associated files (documentation, datasets, etc)

# Function consultation card

## Example 1

The screenshot shows a web interface titled "R Calls" with a tree view of folders. The tree structure is as follows:

- Vera.Georgescu\_netmanuel.R
  - locator
- Vera.Georgescu\_kinetics30h.R
  - tk\_select.list
  - lines
  - lm
  - layout
- tk\_select.list
- lm
- layout

Two callout boxes are present:

- A light blue box with the text "→ Visualization of the call graph" has a yellow arrow pointing to the "locator" folder.
- A light blue box with the text "→ Hypertext links towards semantically related functions" has a yellow arrow pointing to the "Vera.Georgescu\_kinetics30h.R" folder.

## Example 2

The screenshot shows a web interface titled "R Calls" with a tree view of folders:

- abline
- plot

Below the "R Calls" section is a "Sequence" section with the text:

Could be used after: [Juliette.Fabre\\_LocusDiff.R](#)

Two callout boxes are present:

- A light blue box with the text "→ Visualization of the call graph" has a yellow arrow pointing to the "abline" folder.
- A light blue box with the text "→ Hypertext links towards semantically related functions" has an orange arrow pointing to the "Juliette.Fabre\_LocusDiff.R" link.



# Powerful search tools

Build a *SPARQL* request adding conditions on the properties



- Basic Search (if you know the name of the function)  
 Advanced Search

isDedicatedTo

hasIntention

```
PREFIX OntologyR: select ?fonction ?description where { ?fonction OntologyR:isDedicatedTo OntologyR:Phenodyn
OntologyR:hasIntention OntologyR:Visualisation ?fonction OntologyR:hasDescription ?description}
```



There are 4 functions matching your request:

Name	Description
<a href="#">Vera.Georgescu_LERvalidation.R</a>	The general function of visualisation, automatic and manual correction of Elongation Rate kinetics measured on the Phenodyn platform. This function runs on R version 2.6.2
<a href="#">Vera.Georgescu_kinetics30h.R</a>	The function gives a representation of Phenodyn leaf elongation rate kinetics for one night and the following day and night (about 30 h). It performs and represents simple regressions on the nights. It allows manually the LER data and performs the new regressions when data have been invalidated. It displays a selection list that proposes to correct the data, see the following day or come back to the previous day. This function runs on R version 2.6.2

## Example: search the functions

→ Dedicated to the Information System 'Phenodyn' and with an intention of Visualisation

→ That could be used after 'ImportData.R'

→ That call the R core function 'anova'

## Prospects

- Add formal relations with reports, articles, etc
- Perform more automatic extraction from R function documentation

## Prospects

- Add formal relations with reports, articles, etc
- Perform more automatic extraction from R function documentation

## Conclusions

- Users find this repository relevant (efficient search, easy annotating)
- Semantic Web tools allow reasoning for an 'intelligent' repository
- Models and softwares tools are easy to adapt:
  - ↳ for other fields of research
  - ↳ for other programming languages
  - ↳ for mathematical models