



HAL
open science

Bibliométrie, Contrôle qualité et production d'indicateurs. Automatisation sous R

Sylvain Cariou, Lise Frappier

► **To cite this version:**

Sylvain Cariou, Lise Frappier. Bibliométrie, Contrôle qualité et production d'indicateurs. Automatisation sous R. 7. Rencontres R, AGROCAMPUS OUEST. FRA., Jul 2018, Rennes, France. hal-01832284

HAL Id: hal-01832284

<https://hal.science/hal-01832284>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



7èmes Rencontres R - Juillet 2018 - Rennes

Bibliométrie, Contrôle qualité et production d'indicateurs : automatisation sous R

S. Cariou, L. Frappier
SMART-LERECO, INRA, AGROCAMPUS OUEST, 35000, Rennes, France



Qui sommes-nous ?

**SMART
LERECO**

Structures et marchés agricoles, ressources et territoires
Laboratoire d'études et de recherche en économie



**Information Scientifique et Technique (IST) /
documentaliste**



**Gestion et administration de systèmes de gestion de
données**

ist@inra

Contexte

Objectifs

- Contrôle qualité des données
- Production d'indicateurs

Pourquoi

- Gain de temps
- Reproductibilité

Pour qui

- Unités du département SAE2 (19)
 - . Documentalistes // Contrôle qualité
 - . Directions // Note indicateurs

Contexte

Objectifs

- Contrôle qualité des données
- Production d'indicateurs

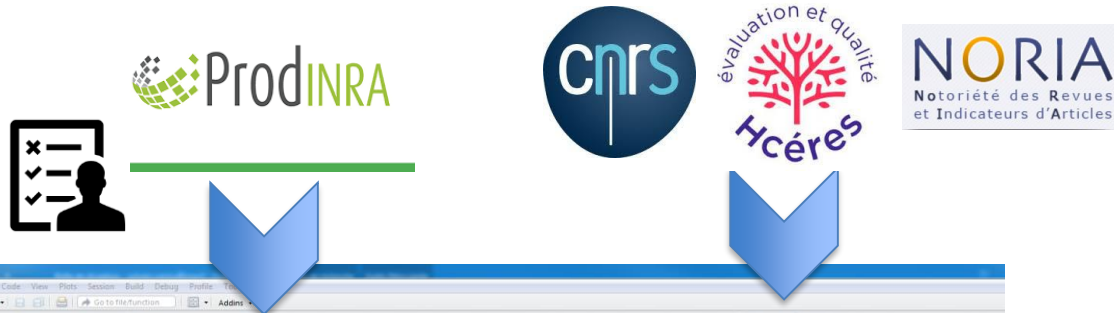
Pourquoi

- Gain de temps
- Reproductibilité

Pour qui

- Unités du département SAE2 (19)
 - . Documentalistes // Contrôle qualité
 - . Directions // Note indicateurs

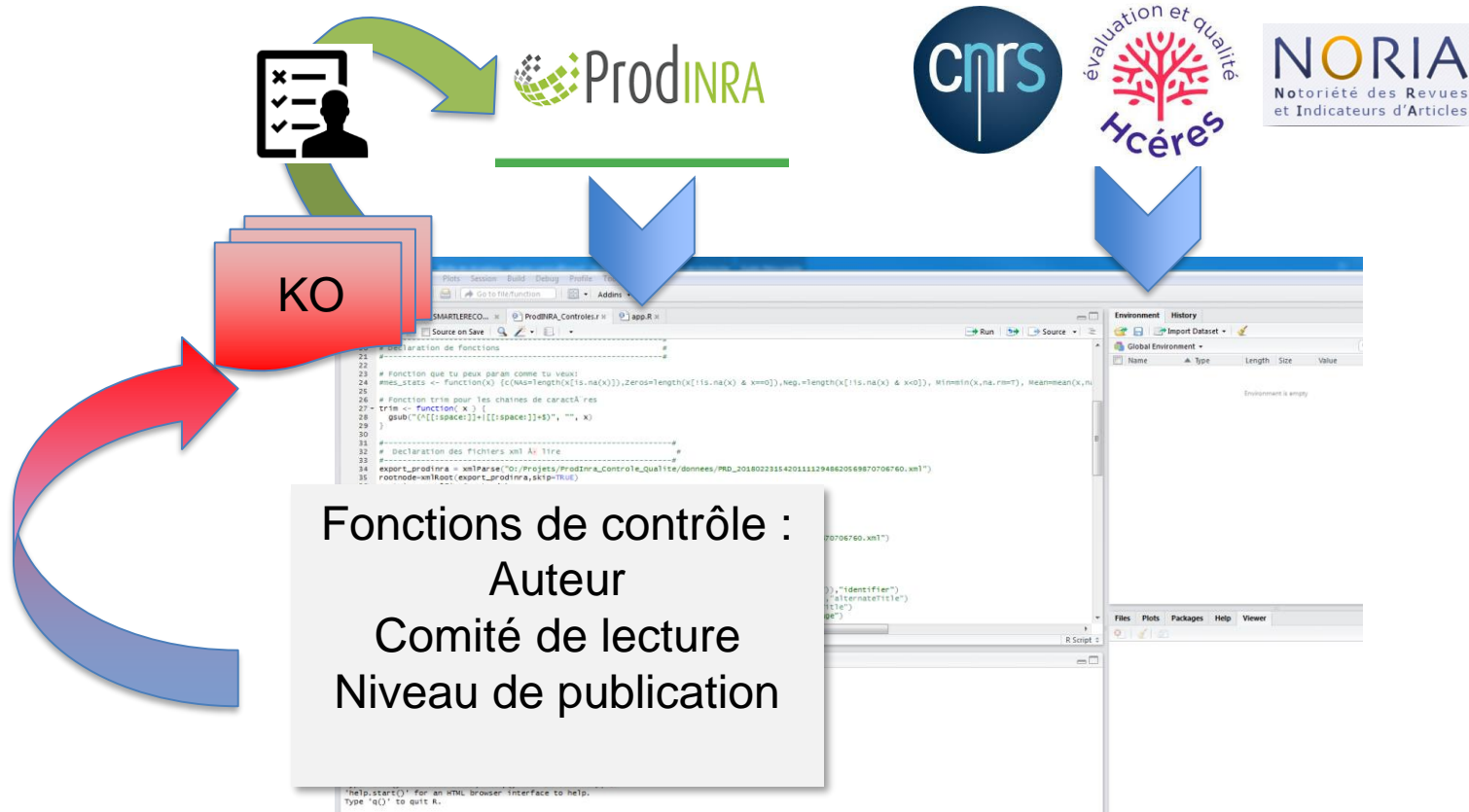
Organisation envisagée de la solution



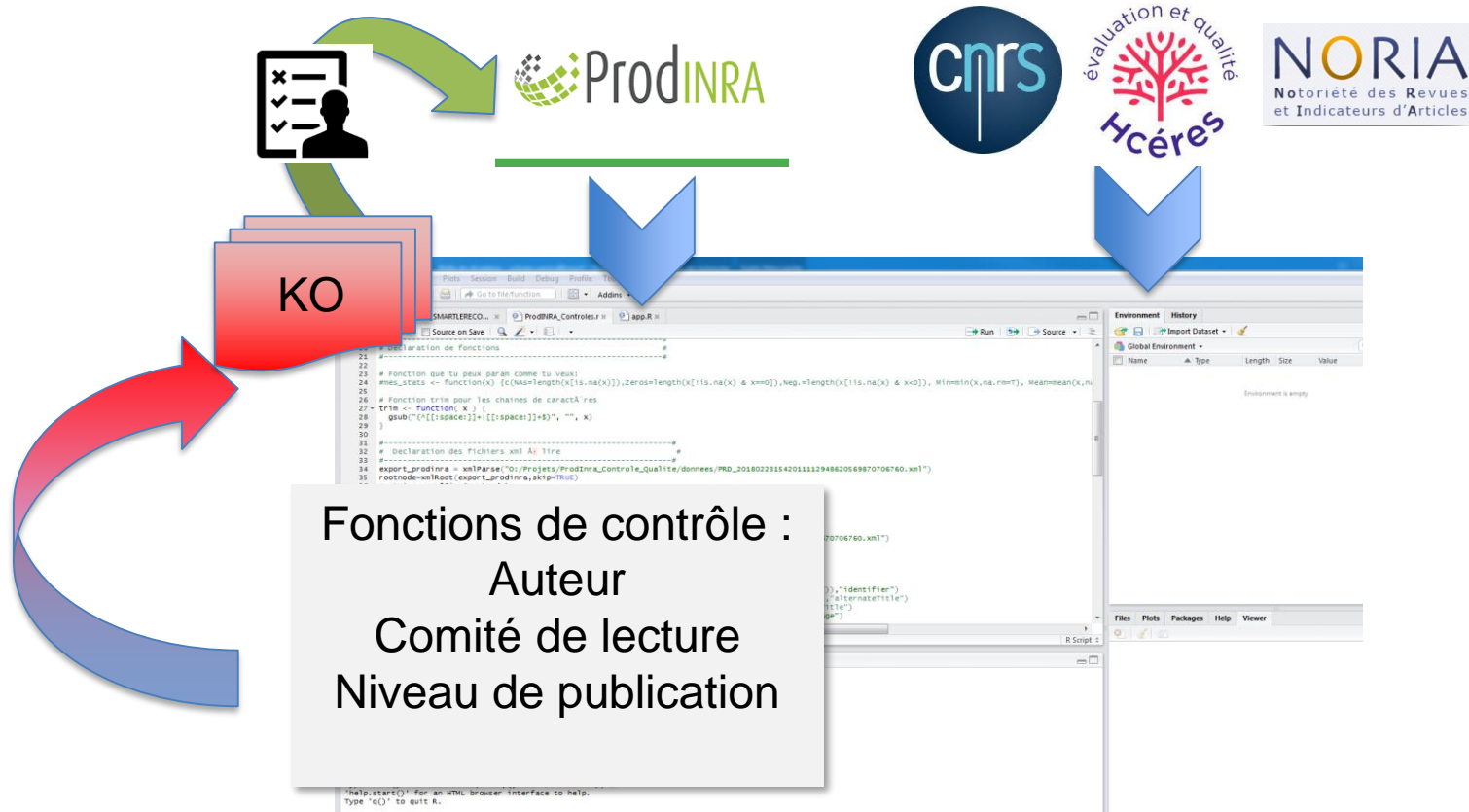
```
20 # déclaration de Fonctions
21 #-----
22 # Fonction que tu peux parain comme tu veux!
23 #mes_stats <- function(x) {c(NAS=length(x[is.na(x)]),zeros=length(x[is.na(x) & x==0]),neg=length(x[is.na(x) & x<0]), min=min(x,na.rm=T), Mean=mean(x,na.rm=T))}
24
25 # Fonction trim pour les chaines de caractères
26 trim <- function(x) {
27   gsub("[[:space:]]+|[[[:space:]]+}$", "", x)
28 }
29
30 #-----
31 # déclaration des fichiers xml À titre
32 #-----
33 export_prodinra <- xmlParse("D:/Projets/ProdINRA Controle Qualité/donnees/PRD_20180223154201112948620569870706760.xml")
34 rootnode <- xmlRoot(export_prodinra, skip=TRUE)
35 rootsize <- xmlSize(rootnode)
36 print(rootsize)
37 print(rootnode)
38 print(rootnode[5])
39 print(rootnode[5])
40
41 test.xmltreeparse("D:/Projets/ProdINRA Controle Qualité/donnees/PRD_20180223154201112948620569870706760.xml")
42 print(test)
43
44
45 dfo.xmlTODataFrame(node=getNodeSet(export_prodinra,"//record/identifier"))
46
47 df_typePub11<-cbind(setNames(xmlTODataFrame(node=getNodeSet(export_prodinra,"//record/identifier")), "identifier")
48   , setNames(xmlTODataFrame(node=getNodeSet(export_prodinra,"//record/alternatetitle")), "alternatetite")
49   , setNames(xmlTODataFrame(node=getNodeSet(export_prodinra,"//record/shorttitle")), "shorttitle")
50   , setNames(xmlTODataFrame(node=getNodeSet(export_prodinra,"//record/language")), "language")
51 )
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
```

BibliometRics ?
Bibliometrix
BibeR

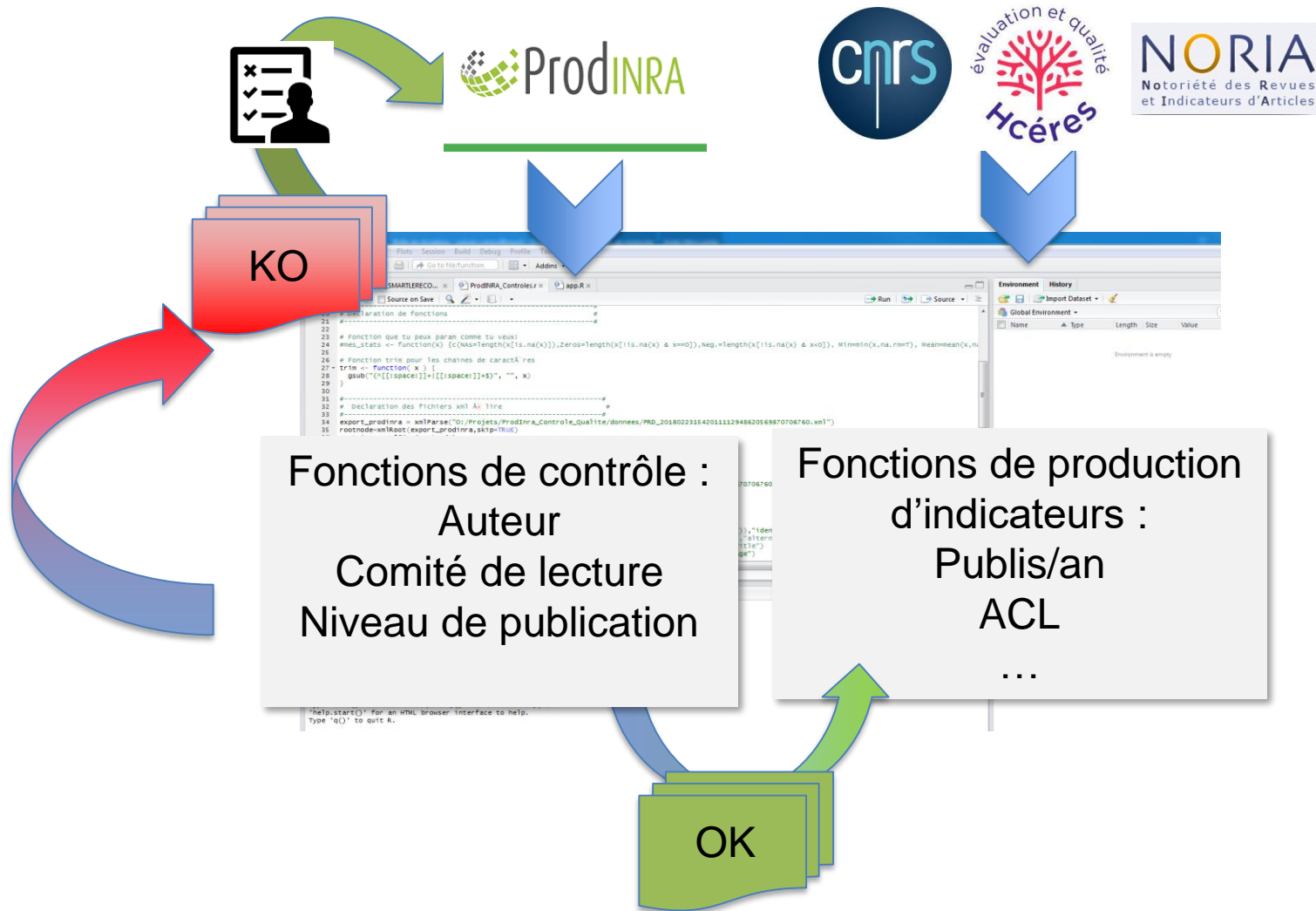
Organisation envisagée de la solution



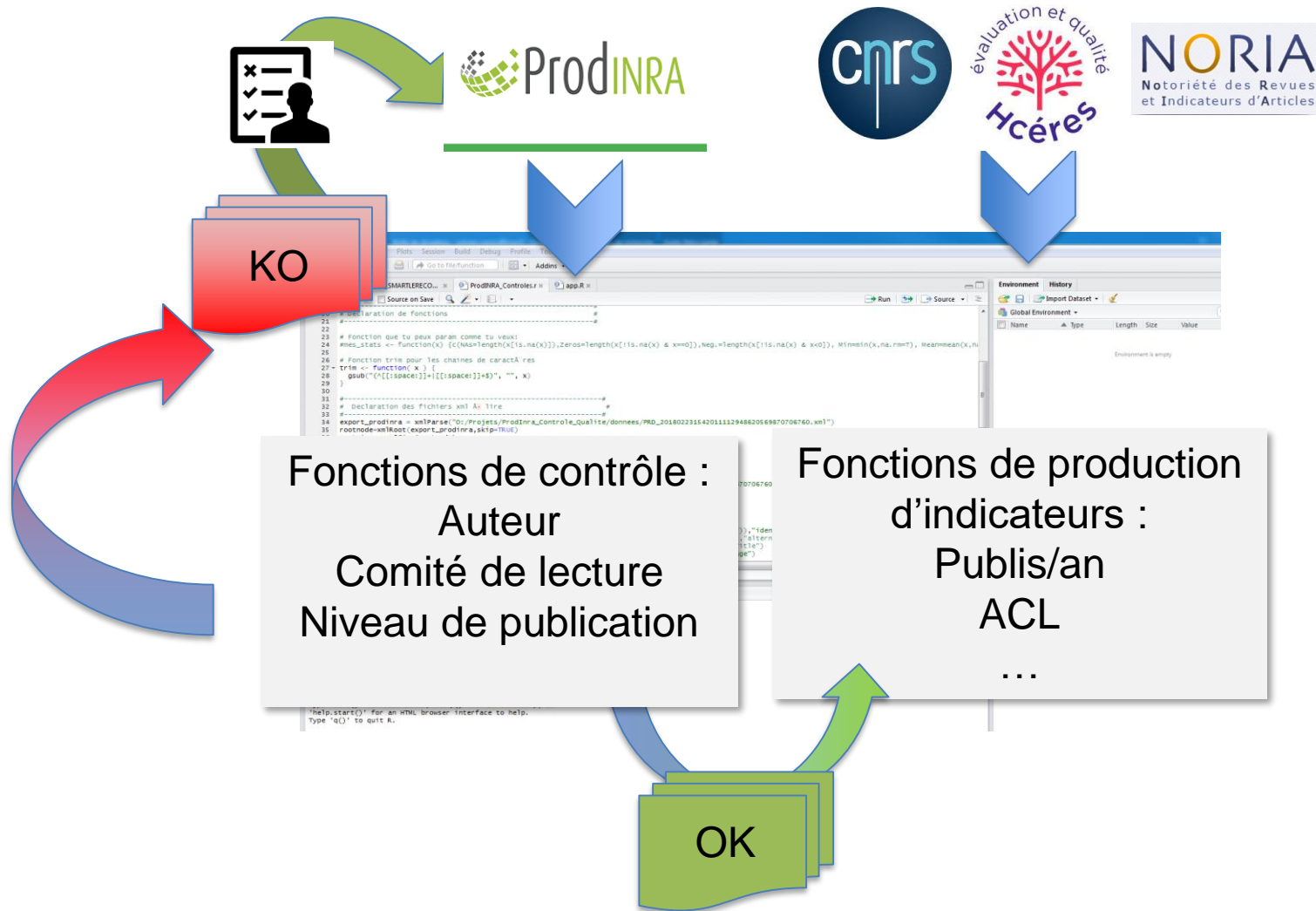
Organisation envisagée de la solution



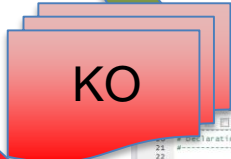
Organisation envisagée de la solution



Organisation envisagée de la solution

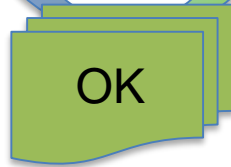


Organisation envisagée de la solution



Fonctions de contrôle :
Auteur
Comité de lecture
Niveau de publication

Fonctions de production
d'indicateurs :
Publis/an
ACL
...



help.start() for an HTML browser interface to help.
Type 'q()' to quit R.

My Documentation
Sat 14 Jul 2017
September 02, 2017

Usage
This is an example of code documentation.

Code Example
Then read in the data.

```
data = read.csv("C:/Users/Start/Desktop/data.csv")
# view data
head(data)
##   product customer_id product_id value
## 1 1 1 1 1 10
## 2 1 1 1 2 20
## 3 1 1 1 3 30
## 4 1 1 1 4 40
## 5 1 1 1 5 10
## 6 1 1 1 6 5
```

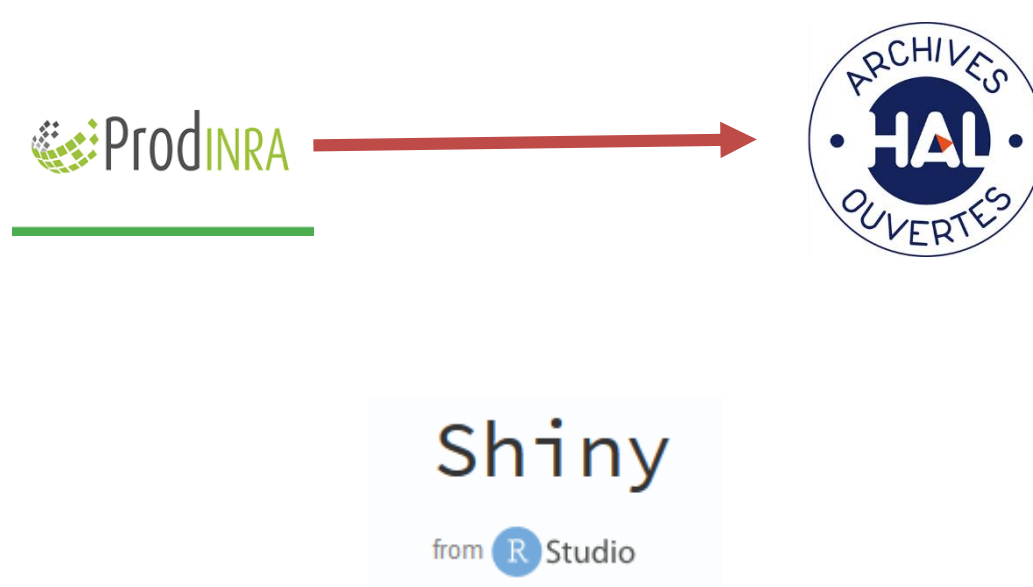
table(data\$product_id)

```
##
## 1 2 3 4 5 6
## 12 8 12 12 12 8
```

Here is a graph where the code is styled using "utf8" font for output to enable.

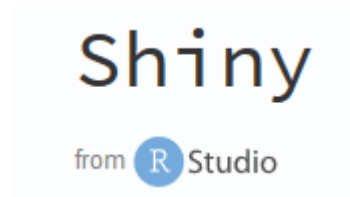
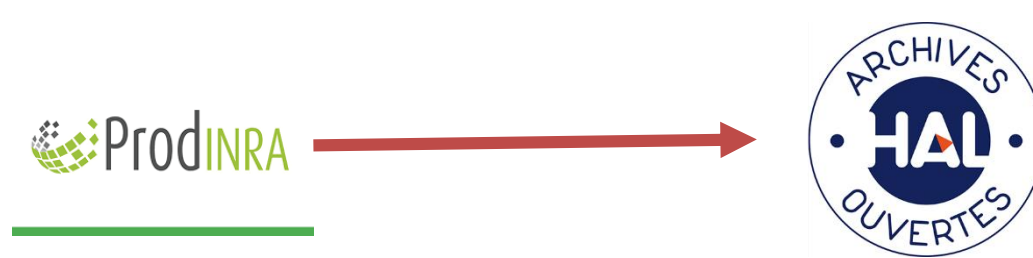
ggplot2
openxlsx

Perspectives et questions



HELP ? Parser xml via R

Perspectives et questions



HELP ? Parser xml via R

Merci de votre attention

Sylvain.Cariou@inra.fr

Lise.Frappier@inra.fr