Trabalho Final - R para Iniciantes

Maurílio Bonora Júnior

26/04/2019

## Introdução

Sempre que vamos escrever um projeto ou relatório, somos bombardeados por um gigantesca base de dados com artigos de relevância variável sobre o determinado assunto. Organizar e escolher os melhores artigos dentre essa imensidão de referências é uma dura tarefa, que muitas vezes pode demandar um tempo que muitos pesquisadores não têm. Contudo, com o advento do R e seus diversos pacotes, entre eles o “Bibliometrix”, essa tarefa de comparar artigos para escolher os melhores se tornou muito mais fácil. Para exemplificar isso, foram escolhidos dois termos para se começar a busca de arquivos em base de dados (Scopus). Tais termos foram: “TCR” e “Signalling”. Abaixo segue todo o procedimento com os códigos e gráficos possíveis de se fazer com o pacote “Bibliometrix”.

## Requerimentos

Para a realização do trabalho, foram necessários alguns programas e pacotes, dentre eles:

* O próprio R: garante toda a linguagem R para se trabalhar;
* RStudio: um software com uma interface mais elegante e amigável ao usuário para ele conseguir trabalho;
* Git: um sistema de controle de versão distribuído. Utilizado juntamente do Rmarkdown para se fazer o upload direto do trabalho para um repositório em nuvem (Github);
* MikTex: programa necessário para a conversão do Script do Rmarkdown para PDf;
* Pacote “Rmarkdown”: converte scripts do R (mais especificamente do Rmarkdown) em uma variedade de formatos incluindo HTML, MS Word, PDF e Beamer. Além disso, o Rmarkdown consegue compilar os scripts em especíes de “livros” onde é possível colocar comentários, códigos fontes e a saída (resultado) do código do script;
* Pacote “Bibliometrix”: garante um conjunto de ferramentas muito útil para análises na área de cientometria e bibliometria;
* Um arquivo Scopus.bib: arquivo baixado da base de dados Scopus com todas as informações de artigos escolhidos a partir das palavras-chaves selecionado.

## Desenvolvimento

### Carragamento e Conversão dos Dados

Após a instalação do R, RStudio, Miktex e Git (lembrando que é necessário a criação de um diretório com o nome do trabalho no Git), é necessário fazer o download dos pacotes “Rmarkdown” e “Bibliometrix”.

Em seguida é necessário carregar o pacote “Bibliometrix”:

library(bibliometrix)

## To cite bibliometrix in publications, please use:  
##   
## Aria, M. & Cuccurullo, C. (2017) bibliometrix: An R-tool for comprehensive science mapping analysis, Journal of Informetrics, 11(4), pp 959-975, Elsevier.  
##   
##   
## http:\\www.bibliometrix.org  
##   
##   
## To start with the shiny web-interface, please digit:  
## biblioshiny()

Já com o arquivo scopus.bib dentro da sua pasta de trabalho está na hora de começar a preparar os dados para as análises. A primeira função que será utilizada é “read.Files”, que converte todos os arquivos de texto de scopus.bib em um grande vetor de caracteres que chamaremos de D.

D<-readFiles("scopus.bib")

Esse grande objeto D pode então ser convertido em um DataFrame a partir da função “convert2df”.

M <- convert2df(D, dbsource = "scopus", format = "bibtex")

##   
## Converting your scopus collection into a bibliographic dataframe  
##   
## Articles extracted 100   
## Articles extracted 200   
## Articles extracted 300   
## Articles extracted 400   
## Articles extracted 500   
## Articles extracted 511   
## Done!  
##   
##   
## Generating affiliation field tag AU\_UN from C1: Done!

### Análise Bibliométrica

O primeiro passo para fazer uma análise descritiva desse Data Frame é utilizara função “biblioAnalysis”, que irá calcular as principais medidas bibliométricas de toda nossa base de dados.

results<-biblioAnalysis(M, sep = ";")

Tal função retorna um objeto da classe bibliometrix (introduzida com o pacote homônimo), que contem os seguintes componentes como o número total de artigos, primeiro autor de cada manuscrito, número de vezes que cada manustrico foi citado, entre outros. Para mais informações use a função:

help("biblioAnalysis")

## starting httpd help server ... done



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.