

Presentación Displex

En este documento se presenta una librería de herramientas en R para el cálculo de la disponibilidad léxica, para la que se pretende que sea una guía rápida de uso básico. Aunque la librería permite, a través de programación funcional, implementar múltiples modelos, se proporcionan un conjunto de herramientas básicas prediseñadas para que se pueda empezar a usar de forma casi inmediata. Sin embargo, consideramos necesario hacer notar que estas funciones de aplicación directa representan a un subconjunto de las posibilidades que permite el sistema y que se han utilizado como herramienta rápida más que como solución definitiva.

Hay otro trabajo en preparación en el que se discuten en profundidad los modelos y diferentes opciones, así como un estudio en profundidad de diferentes aspectos del cálculo de la disponibilidad y su interpretación. Sin embargo, el desarrollo de estas herramientas posteriores requeriría cierta soltura en el manejo de la sintaxis de R, así como con los conceptos de programación funcional.

Instalación

Se ha optado por el uso del repositorio GitHub.com. Esto implica que el usuario debe llevar a cabo un pequeño paso de instalación que, en nuestra opinión, no es más complejo que cualquier proceso de instalación en el sistema R. Este paso ha de hacerse únicamente cuando se pretenda instalar o actualizar el paquete en el sistema. Puesto que nuestra intención es seguir trabajando en el paquete, sería recomendable realizar esta acción de forma periódica.

Las siguientes órdenes instalan el paquete Displex del repositorio GitHub. Todo el código está implementado en R, con lo que es fácil de revisar.

```
install.packages("devtools")
library(devtools)
install_github("jmss70/displex")
```

Se recomienda el uso del universo TidyVerse para el análisis de datos, ya que proporciona herramientas con una sintaxis muy potente que facilita enormemente la tarea de análisis y representación de datos.

```
install.packages("tidyverse")
```

Llevados a cabo los pasos anteriores, y si no se ha producido ningún contratiempo, el sistema está preparado para llevar a cabo el trabajo.

Carga de las librerías y los datos

Para poder usar las librerías hay que cargarlas en nuestra sesión de trabajo, poniendo a nuestra disposición las funciones que proporcionan:

```
library(tidyverse)
```

```
## -- Attaching packages -----
## v ggplot2 3.3.2      v purrr   0.3.4
## v tibble  3.0.3      v dplyr   1.0.2
## v tidyr   1.1.1      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
```

```
## -- Conflicts -----
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()
```

```
library(displex)
```

Para cargar los datos, se espera que estén en un determinado formato, utilizado por ser el que habitualmente se encuentra en este tipo de estudios, creemos que por razones históricas. Sin embargo, consideramos que son redundantes y que habría que estudiar, en un futuro cercano, establecer un estandar de codificación que sea más coherente con los modelos de datos normalizados.

Se espera que los datos estén en un archivo de texto, con campos separados por espacios: - Un campo de información del hablante - Un campo de identificación de usuario - Un campo de identificación de centro de interés - Una lista de palabras separadas por comas y en orden de realización

Un ejemplo de dos líneas sería:

```
21131 001 01 mano, pie, brazo, cerebro, pulmón, nariz, extremidad, ojo, boca, diente, pelo, oreja, culo
12131 002 01 riñón, corazón, garganta, cabeza, pierna, pie, hígado, estómago, mano, brazo, antebrazo, al
```

Suponiendo que tenemos todos los datos cargados en un archivo, denominado `datos.txt`, que estará alojado en el mismo directorio que el script de procesamiento, se podrían cargar los datos como:

```
d <- read.displex("datos.txt")
head(d)
```

```
##   infos users centers      words
## 1 21131   001      01 mano, pi...
## 2 12131   002      01 riñón, c...
## 3 12213   003      01 brazo, m...
## 4 22214   004      01 brazo, o...
## 5 12214   005      01 cabeza, ...
## 6 22213   006      01 pie, man...
```