

Word Cloud from pdf

Esta aplicación tiene como finalidad crear nubes de palabras a partir de ficheros pdf. Esta centrada particularmente en crear nubes de palabras sobre asignaturas relacionadas con informática. A partir de las fichas docentes de una asignatura es capaz de generar una nube con las palabras más repetidas de esa asignatura.

Cómo usar la aplicación

Para integrar las palabras que más se repiten de un pdf en una página web html debemos seguir los siguientes pasos:

1. Introducimos en la carpeta **fichas_docentes** todas las fichas docentes de las que se quiera generar una nube. Se generará una nube por cada fichero con extensión pdf. Este es el único formato admitido.
2. Ejecutamos el programa **pdf2htmlWordCloudAll.js** si vamos a generar más de una nube de palabras. En caso de que solo queramos generar una única nube ejecutamos el programa **pdf2htmlWordCloud.js** con el nombre del fichero pdf del cual queremos extraer la nube (sin extensión), por ejemplo: `node pdf2htmlWordCloud ficha_docente`.
3. En la carpeta **nubes_ficha_docente** se encuentran todas las páginas html generadas con las nubes de las palabras que más se repiten en cada pdf.
4. Estas nubes se pueden integrar en otras páginas. La página html **complutense.html** es un ejemplo de cómo integrar las nubes a la página principal de fichas docentes de la Universidad Complutense de Madrid.