

## **RECUPERACIÓN DE LA INFORMACIÓN**

**21 de noviembre de 2019**

### **PRÁCTICA 3 – Recuperación de la información con el modelo vectorial**

Usando el código generado en la primera práctica se deberá diseñar un programa para recuperación de la información basado en el modelo vectorial. Las características del sistema son las siguientes:

- a) Cada documento es un fichero de texto puro con extensión .txt
- b) Hay un fichero de "palabras vacías" que deberá usarse para preprocesar cada documento con el fin de obtener su representación normalizada.
- c) Cada documento es representado mediante una serie de términos en un fichero de texto de igual nombre con extensión .rep
- d) El programa ejecutable a desarrollar tendrá opciones para:
  - 1. Generar o actualizar el fichero .rep de representación de uno o todos los documentos en base a su contenido y el fichero de palabras vacías.
  - 2. Realizar una consulta usando el modelo vectorial mediante una lista de términos para entregar como resultado cero, uno o varios documentos.
  - 3. Salir

Se deberán entregar todos los ficheros utilizados, incluyendo el programa fuente, y un informe en formato PDF sobre los criterios de diseño e implementación de cada aspecto del desarrollo de la práctica. Consultar al profesor sobre el formato del informe.

**Fecha de entrega 5 de diciembre de 2019.**