Diagrama UML del sistema:

Proceso para la recuperación de información correpondiente para los modelos vectorial, booleano y probabilístico:

* Lista (lista)
* cuentaPal (lista)
* proba (double)
* Diccionario()(Dict)
* calcProb()(double)
* resultado()(Dict)

Modelo Probabilisticoado

* listaPal (lista)
* Dic\_Pal (Dict)
* Diccionario()(Dict)
* matrizInc()(Dict)
* procCon()(Dict)

Modelo Booleanoado

* listaPalabra (lista)
* cuentaPal (lista)
* frecuencia (int)
* vector (lista)

Modelo Vectorialado

* Archivo (txt)
* Ar\_tockenizado (lista)
* Doc\_etiquetado (lista)
* lecturaArchivo() (lista)
* tockenizacion() (lista)
* etiquetado() (lista)
* quitarStopWords() (lista)

Proceso de Indexadoado

Los tres modelos presentan sus resultados de mayor relevancia a menor relevancia, acabando con la fase de recuperación de la información.

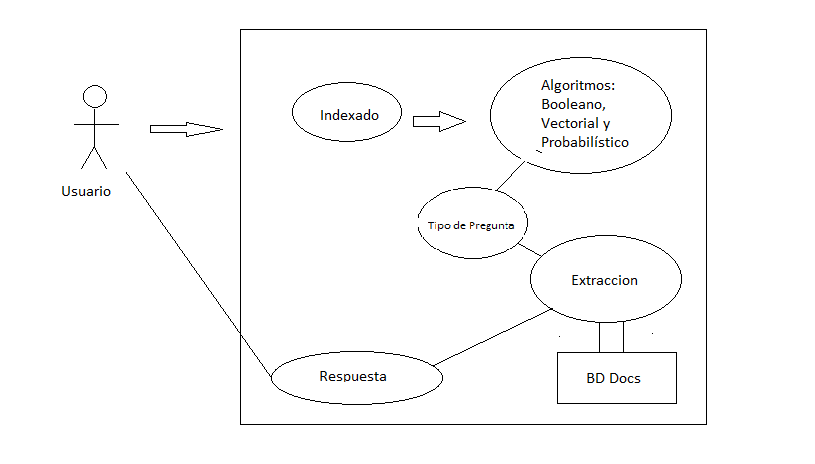
* cuentaPal() (lista)
* tfidf() (lista)
* normalizacion()(lista)
* resultado() (lista)

Procesamiento de la pregunta

* listaPreguntas (lista)
* listaTipoPregunta (lista)
* preguntaUsuario (String)
* registraPregunta() (lista)
* regexchunk() (lista)
* comparaPregunta() (Dict)
* resultado() (lista)

En la fase de extracción es importante conocer el tipo de pregunta para accesar al vocabulario de búsqueda de respuesta.

Casos de uso:



El usuario ingresa cadenas para formular la consulta con las preguntas ¿Qué? ¿Quién? ¿Cuándo? ¿Dónde? Y el sistema recuperará una lista documentos junto con el valor porcentual de mayor a menor importancia.