

Universidad Nacional de Mar del Plata

Facultad de ingeniería

Carrera : Ing. Informática – Cátedra organización de datos 2016

TP integrador índices invertidos y motores de búsqueda

Realice los 4 pasos

- 1) Parseo
- 2) Ordenamiento
- 3) Agrupamiento.
- 4) Generación del Diccionario y Posting.

Para los siguientes textos

“Ya no me necesitas es lo mejor
Eras alguien a quien yo solía conocer
Fue muy simple despegar
Solo un corto tiempo
Y te buscaste un nuevo corazón

Ahora tienes tu propio show
Como un rey vengador, vengador
No te alcanza con improvisar
El descaró baby es parte de la diversión”

“¿Que importan ya tus ideales?
¿Que importa tu canción?
La grasa de las capitales
cubre tu corazón.
¿Por qué tenes que llorar?
Es que hay otro en tu lugar que dice:
"Vamos, vamos, la fama."
Tu oportunidad está ahí,
lo mismo me pasó a mí.
Lo tienes todo, todo
no hay nada.”

Ejercicio de programación

Implemente una aplicación con la siguiente funcionalidad general:

Debe procesar un directorio (tomado como parámetro) y del mismo indexar todos los archivos de texto que encuentre en él, permitiendo posteriormente realizar búsquedas por palabras y obtener la lista de archivos donde aparecen estas palabras. Para la indexación debe generar las estructuras necesarias (pueden ser solo en memoria), utilizando índices inversos. Como mínimo permitirá una búsqueda por una sola palabra. Se realizará el control mínimo de errores (path incorrecto o mal formateado, ausencia de archivos, etc).

Universidad Nacional de Mar del Plata

Facultad de ingeniería

Carrera : Ing. Informática - Cátedra organización de datos 2016

Se debe presentar código fuente, binario ejecutable (explicando la plataforma en que corre) y archivos de prueba. El diccionario debe quedar en un archivo de texto luego de cada ejecución del programa.

Ejemplo de ejecución:

```
hernan@hal9000:/home/sources/indx$./indx ./pruebas
```

```
3 archivo(s) encontrado(s)
```

```
Parseando ... 100% listo
```

```
Ordenando ... 100% listo
```

```
Agrupando ... 100% listo
```

```
Generando diccionario y posteando ... 100% listo
```

```
Ingrese la palabra a buscar : GATO
```

```
Encontrada en el archivo : texto1.txt , 3 veces
```

```
Encontrada en el archivo : texto4txt , 1 veces
```

```
Buscamos otra palabra? NO
```

```
hernan@hal9000:/home/sources/indx$
```