

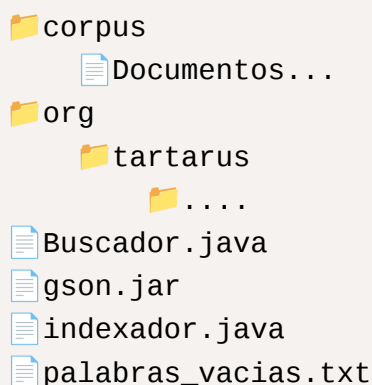
# Manual de Uso

*Alejandro Díaz Gómez y Juan Luis Sena Cárdenas*

Es necesario mencionar previamente que todo el sistema ha sido probado con la versión de java 19 (build 19+36-2238), no se asegura en ningún momento el funcionamiento del mismo en otras versiones.

## Estructura de archivos

El sistema tiene una serie de dependencias obligatorias para su correcto funcionamiento, las cuales son ya incluidas en el código fuente del mismo. Por defecto, la estructura de archivos necesaria es la siguiente, aunque esta puede ser cambiada mediante los parámetros de cada uno de los archivos java:



```
corpus
├── Documentos...
├── org
│   ├── tartarus
│   └── ....
├── Buscador.java
├── gson.jar
├── indexador.java
└── palabras_vacias.txt
```

## Indexador

Para compilar y ejecutar el indexador es necesario establecer previamente los parámetros necesarios. Estos se pueden modificar en las primeras líneas del archivo *indexador.java*, tras la correcta configuración, las instrucciones a introducir en terminal son:

Compilar: `javac -cp . indexador.java`

Ejecutar: `java -cp . indexador`

Tras esto, el módulo indexará los documentos y extraerá los resultados en los archivos con los nombres indicados en los parámetros correspondientes, mostrando por terminal el tiempo requerido para comprobar la operación así como el número de términos indexados.

Para comprobar el correcto funcionamiento del indexador, se ha facilitado una carpeta (*test*) con únicamente 4 documentos y un total de 14 términos para de este modo poder asegurar que los cálculos se realizan correctamente

## Buscador

El proceso de compilación y ejecución del módulo de búsqueda es bastante similar al indexador. Nuevamente hay que establecer los parámetros necesarios en las primeras líneas del archivo *Buscador.java*. La diferencia esta vez es que hay que especificar el uso de la librería gson en la compilación y ejecución. Los comandos son los siguientes:

Compilar: `javac -cp .;gson.jar indexador.java`

Ejecutar: `java -cp .;gson.jar indexador`

Al ejecutar el módulo, se pedirá inmediatamente al usuario que realice una consulta. Las consultas pueden ser de dos tipos:

- **Búsqueda de términos:** El usuario podrá introducir uno o varios términos, en el segundo caso también se dará la opción de escoger entre la aplicación de conjunción (AND), donde los documentos deberán tener relación con todos los términos; o disyunción (OR), donde solo es necesario que un documento de la consulta tenga relación con un documento para que este se incluya en la respuesta.

En este tipo de consulta la respuesta será únicamente la lista de documentos ordenados por relevancia de mayor a menor.

Un ejemplo de uso puede ser el siguiente:

```
Búsqueda: Milky Way
Tipo: AND
Resultado: [000084578400037, 000294852400024, ....., 00023845280
```

- **Búsqueda de frases:** Cuando el usuario introduce una frase entre comillas, el sistema devolverá únicamente los documentos donde aparecen esas frases.

La salida de este tipo de consulta es algo diferente a la búsqueda de términos, en este caso se devuelven aquellos documentos que contengan una o más ocurrencias de la frase. Los nombres de estos serán mostrados junto con los índices del primer y el último término de la consulta para cada ocurrencia.

Un ejemplo más claro de lo que se devolvería para este caso es el siguiente:

Consideremos el documento D1 cuyo contenido se muestra a continuación:

```
D1: The formation and early evolution of the Milky Way Galaxy
Búsqueda: "Milky Way Galaxy"
Resultado: D1: [ [7, 9] ]
```