
Nombre del Sistema:

KoolFileIndexer

Objetivo General del Sistema:

Indexar archivos de un equipo local, extrayendo y almacenando sus metadatos en una base de datos relacional, para facilitar su búsqueda, organización y clasificación mediante filtros, palabras clave y categorías automáticas y generadas por el usuario.

Requerimientos funcionales

- RF_1** El **usuario** debe ser capaz de buscar entre todos los archivos a los cuales el sistema tenga acceso
- RF_2** El **usuario** debe ser capaz de refinar la búsqueda con palabras claves asociadas.
- RF_3** El **usuario** debe ser capaz de refinar la búsqueda con el nombre del archivo.
- RF_4** El **usuario** debe ser capaz de refinar la búsqueda con la extensión del archivo.
- RF_5** El **usuario** debe ser capaz de refinar la búsqueda con el tamaño del archivo.
- RF_6** El **usuario** debe ser capaz de refinar la búsqueda por criterios de búsqueda personalizados (A partir de ahora: etiquetas).
- RF_7** El **usuario** debe ser capaz de refinar la búsqueda por categorías.
- RF_8** El **usuario** debe ser capaz de crear categorías para organizar archivos.
- RF_9** El **usuario** debe ser capaz de definir criterios para cada categoría (palabra clave incluida, fecha de creación en rango, que el nombre contenga un término específico)
- RF_10** El **usuario** debe ser capaz de editar el nombre de las categorías.
- RF_11** El **usuario** debe ser capaz de editar los criterios de las categorías.
- RF_12** El **usuario** debe ser capaz de eliminar las categorías.
- RF_13** El **usuario** debe ser capaz de asociar manualmente etiquetas a los archivos, creándolas si no existen.
- RF_14** El **usuario** debe ser capaz de modificar el nombre de las etiquetas presentes.
- RF_15** El **usuario** debe ser capaz de eliminar las etiquetas existentes
- RF_16** El **usuario** debe ser capaz de asociar manualmente palabras clave a los archivos.
- RF_17** El **usuario** debe ser capaz de modificar las palabras clave de los archivos.
- RF_18** El **usuario** debe ser capaz de eliminar las palabras clave de los archivos.
- RF_19** El **usuario** debe ser capaz de excluir carpetas específicas de ser indexadas.

- RF_20** El **usuario** debe ser capaz de excluir archivos específicos de ser indexados.
- RF_21** El **sistema** debe ser capaz de leer el contenido interno de ciertos tipos de archivo (TXT, PDF, DOCX, archivos de texto simple) y obtener palabras clave del contenido.
- RF_22** El **sistema** debe eliminar las etiquetas cuando no haya archivos asociados a ellas.
- RF_23** El **sistema** debe extraer los metadatos de los archivos para añadirlos al índice de archivos.
- RF_24** El **sistema** debe actualizar automáticamente el índice de archivos si se añade un nuevo archivo.
- RF_25** El **sistema** debe actualizar automáticamente el índice de archivos si se modifica uno de los archivos registrados en este.
- RF_26** El **sistema** debe actualizar automáticamente el índice de archivos si se elimina uno de los archivos registrados en este.
- RF_27** El **sistema** debe añadir al índice de archivos las palabras clave de cada archivo.
- RF_28** El **sistema** debe indexar todos los archivos a los que tenga acceso, siempre que el usuario no especifique lo contrario (ver **RF_19** y **RF_20**).
- RF_29** El **sistema** debe clasificar automáticamente los archivos en las categorías que tengan criterios definidos.
- RF_30** El **sistema** debe generar categorías de forma automática a partir de la extensión y la ubicación del archivo.
- RF_31** El **sistema** debe validar entradas de texto en campos de usuario, como etiquetas, palabras clave y categorías, garantizando que no estén vacías, contengan caracteres válidos y no sean duplicadas.
- RF_32** El **usuario** debe ser capaz de asociar manualmente categorías a los archivos.

Requerimientos No funcionales

- RNF_1** El **sistema** debe proporcionar una interfaz gráfica con retroalimentación visual inmediata, disposición consistente de los elementos en pantalla, que tenga una curva de aprendizaje menor a 10 minutos para tareas básicas como indexar archivos, crear categorías y etiquetar archivos, implementada con JavaFX.
- RNF_2** El **sistema** debe ejecutarse correctamente en al menos dos sistemas operativos: Windows y Linux.
- RNF_3** El **sistema** debe soportar la indexación de directorios con *10.000 o menos* archivos sin una pérdida considerable de rendimiento.
- RNF_4** El **sistema** debe realizar búsquedas en menos de 2 segundos sobre una base de datos con más de 1.000 archivos indexados.
- RNF_5** El **sistema** debe ejecutar la indexación en segundo plano sin afectar la interacción del usuario con la interfaz.
- RNF_6** El **sistema** debe almacenar los índices de los archivos en una base de datos PostgreSQL.
- RNF_7** El **sistema** debe manejar errores sin interrumpir la ejecución, continuando el proceso de indexación incluso si encuentra archivos corruptos o inaccesibles.

Reglas de Negocio Preliminares:

RNP_1 No se deben indexar carpetas del sistema o protegidas (por ejemplo C:/Windows).

RNP_2 Cada archivo puede tener múltiples palabras clave, pero ninguna de ellas debe repetirse dentro del mismo archivo.

RNP_3 Cada archivo puede tener varias etiquetas.

RNP_4 Si un archivo ya está indexado y no ha cambiado, debe evitarse una nueva inserción.

RNP_5 No se indexarán archivos .exe o .dll (archivos ejecutables o en binario)

RNP_6 No todos los archivos deben tener una categoría, una etiqueta o palabras claves, por lo tanto, si se elimina una categoría, no es necesario que los archivos en esta se reasignen a otras categorías.

RNP_7 Las etiquetas y las categorías deben tener nombres únicos.

RNP_8 Se debe priorizar la indexación de los archivos más recientes.

RNP_9 El índice debe contener al menos: nombre, ruta, tamaño, fecha de modificación y extensión.

RNP_10 Las búsquedas se realizan exclusivamente sobre los datos del índice en la base de datos, no directamente sobre el disco en tiempo real.

RNP_11 Cada archivo puede tener máximo una sola categoría.

RNP_12 Las entradas de texto solo pueden contener letras (a–z, A–Z), números (0–9), espacio simple entre palabras (no se permiten espacios múltiples), guion medio (-) y guion bajo (_). No se permiten caracteres especiales como @, #, \$, %, &, /, ?, ni emojis. No se permite que la entrada sea vacía ni solo espacios en blanco.

RNP_13 Una carpeta se considera **inválida** si:

- Ha sido eliminada entre la selección y la validación.
- Está vacía o no contiene archivos válidos para indexación
- Se encuentra bloqueada por el sistema operativo o protegida por políticas de seguridad.

RNP_14 Una etiqueta no puede existir si no tiene archivos asociados.

RNP_15 Una categoría sí puede existir incluso sin archivos asociados.

RNP_16 Si un archivo cumple todos los criterios de una categoría es asignado a esta, si cumple todos los criterios de varias, se asigna a todas esas categorías.

RNP_17 Las **rutas de archivo** se almacenan siempre en formato absoluto y normalizado, sin segmentos “. .” ni enlaces simbólicos, y las comparaciones se realizan de forma *case-insensitive*.

RNP_18 Los **nombres de archivo** se guardan con sus espacios al inicio o fin eliminados, pero conservan mayúsculas/minúsculas.

RNP_19 Solo existen categorías **automáticas** (por extensión o ubicación); el usuario no crea categorías manuales en esta versión.

RNP_20 Un archivo con extensión o ubicación no reconocida recibirá la categoría genérica “**Sin categoría**” (no se permite `null`).

RNP_21 **Longitud máxima uniforme** para simplificar y homogeneizar el control de entradas, de modo que todas las etiquetas, categorías y palabras clave deben tener un mínimo de un carácter significativo y un máximo de cincuenta caracteres, lo cual evita inconsistencias de usabilidad y mantiene la interfaz predecible.

RNP_22 Las **palabras clave** no pueden contener espacios; deben ser un token continuo de 1–50 caracteres válidos (a–z, 0–9, _, -). Para conceptos de varias palabras, usar **etiquetas**.

ALCANCE Y RESTRICCIONES

Alcance del sistema:

- El sistema KoolFileIndexer será una aplicación de escritorio local que permite al usuario indexar archivos de su equipo, extraer sus metadatos, y almacenarlos en una base de datos PostgreSQL para realizar búsquedas eficientes y organizar los archivos.
- El sistema permitirá búsquedas por diferentes criterios, como nombre, tamaño, extensión, palabras clave, etiquetas y categorías definidas por el usuario.
- Se podrá crear, modificar y eliminar etiquetas y categorías. Las etiquetas serán asignadas directamente por el usuario, mientras que las categorías podrán ser asignadas manual o automáticamente según criterios definidos.
- El sistema incluirá mecanismos de validación de entrada de texto para etiquetas, palabras clave y nombres de categorías, garantizando que no haya duplicados, campos vacíos ni caracteres inválidos.
- La aplicación clasificará automáticamente los archivos al momento de indexarlos o modificarlos, aplicando los criterios definidos por el usuario (palabra clave incluida, un término específico en el nombre o fecha de creación).
- El sistema actualizará el índice automáticamente cuando detecte archivos nuevos, modificados o eliminados.

Restricciones del sistema:

- El sistema funcionará únicamente en entorno local, sin conexión a servidores web ni microservicios.
- El motor de base de datos debe ser relacional.
- No se podrán indexar carpetas protegidas por el sistema operativo ni aquellas consideradas inválidas (vacías, eliminadas o sin permisos).

- Las etiquetas deben ser únicas por archivo, tener nombre único, y no pueden existir si no están asociadas a al menos un archivo. - pueden tener varias etiquetas
- Las categorías deben tener nombres únicos a nivel de sistema.
- Cada archivo puede tener una sola etiqueta, múltiples palabras clave, y pertenecer a varias categorías definidas por el usuario.
- Cada archivo solo puede pertenecer a una única categoría **automática** (generada por el sistema), pero sí puede estar en múltiples categorías definidas por el usuario.
- Las entradas de texto (etiquetas, palabras clave y categorías) solo pueden contener letras (a–z, A–Z), números (0–9), espacios, guion medio (-) y guión bajo (_). No se permiten caracteres especiales ni entradas vacías o de solo espacios.
- Las búsquedas y la clasificación automática se activan únicamente cuando:
 - Se indexa un nuevo archivo.
 - Se modifica un archivo ya indexado.
- El sistema no debe depender de servicios externos ni bibliotecas que requieran conexión a internet para su funcionamiento principal.