

## Bases de Datos 2024 - Segundo Parcial de muestra

**Ejercicio 1:** Sea la siguiente base de datos de fotografías:

usuarios (uid, nombre)

fotos (fid, dueño, tamaño)

comentarios (cid, autor, foto, texto)

Observar que *dueño* y *autor* son usuarios (*uid*), *foto* es una *fid*.

Considerar la consulta:  $\prod_{fid, uid} (fotos \bowtie \sigma_{nombre='Luis\ Pérez'} (usuario))$ .

Asumir que hay índice primario en *nombre* para usuario, que usuario tiene 100.000 tuplas, que un bloque tiene 4096 bytes, que hay 10.000 nombres diferentes de usuarios, que un mismo nombre no se repite más de 15 veces en usuario.

1. Calcular la cantidad de claves de búsqueda por nodo del índice (árbol B+)
2. Calcular la altura del árbol B+.
3. Indicar el operador físico que conviene usar para la selección y estimar el costo en número de transferencias de bloques para este operador físico.

**Ejercicio 2:** Recuperación de la información: Responder:

1. Supongamos que queremos hacer búsquedas ordenadas por relevancia de documentos en una colección de documentos en mi disco rígido; asumir que se usa modelo de espacio vectorial e índices invertidos. Las consultas son booleanas. Dar significado de las componentes de las tuplas que van en una lista de índice invertido asociada a un término.
2. Listar los pasos para procesar una consulta basada en relevancia de documentos; es una colección que está en un disco rígido, se usa modelo de espacio vectorial, similitud de vectores para la relevancia, y consultas booleanas. Evitar pasos innecesarios.

**Ejercicio 3:** Identificar los distintos tipos de contexto que aparecen en el siguiente texto:

“Mientras tomaba un mate en la plaza principal de mi pueblo, escuchando el sonido de las campanas de la iglesia, recordé las protestas globales por el cambio climático. Pensé en nuestros ancestros, que vivían en armonía con la naturaleza, y en cómo nuestra sociedad actual ha perdido esa conexión. Me pregunté si volveremos a esos tiempos, o si seguiremos repitiendo los mismos errores del pasado.”

**Ejercicio 4:** ¿Hay alguna capa de un codificador de un transformer que haga reconocimiento de intenciones? ¿Alguna capa de decodificador hace la determinación del contenido? ¿Alguna capa del decodificador hace estructuración de la respuesta y planificación de frases? ¿Un transformer hace un post procesamiento de su output?