

FASE 1: IDENTIFICACIÓN DEL PROBLEMA

Definición del problema.

El equipo VIP de Simulación de la universidad ICESI requiere de un prototipo software que le permita gestionar eficientemente un gran flujo de información personal de residentes del continente americano a través de una base de datos que permita las operaciones básicas de la misma (CRUD)

Identificación de necesidades y síntomas.

- El grupo VIP requiere una solución de software que le permita la generación automatizada de una gran cantidad de personas simuladas con el fin de probar el funcionamiento de estructuras de datos eficaces.
- La solución requerida debe brindar una manera eficiente de generar, almacenar y gestionar datos.
- El modelamiento de la solución debe ofrecer características especiales para una futura búsqueda autocompletada de parametrizados.
- La generación de los datos simulados debe tener coherencia y estar acorde con data sets previamente estipulados.

FASE 2: RECOPIACIÓN DE LA INFORMACIÓN NECESARIA

Requerimientos para la solución del problema

- Generar una cantidad específica de personas simuladas.
- Agregar este set de personas simuladas a una base de datos.
- Se debe permitir al usuario crear, editar y buscar personas individuales.
- Se debe desplegar un listado con los posibles resultados de una búsqueda y además permitir la edición de cada uno de estos.

Restricciones para la solución del problema

- Las operaciones crear, leer, actualizar y eliminar datos en la solución deben modelarse bajo un concepto de alta eficiencia.
- Para la búsqueda de datos se debe implementar estructuras de modelamiento de tipo árbol binario de búsqueda auto balanceado (AVL).
- La solución de software debe contar con una interfaz gráfica de usuario y una disposición de sus elementos estipulada previamente.
- Este subproyecto de software debe entregarse junto con sus respectivas pruebas automatizadas en las cuales se evidencia el correcto funcionamiento en su totalidad.
- Implementación de estructuras basadas en el concepto trie.

Definiciones

- **Generación simulada de personas:** Consta de crear perfiles de personas con información ficticia.

- **Data set:** Gran conjunto de información recopilada y de diversas fuentes orientada a aspectos específicos.
- **CRUD:** (Por sus siglas en inglés *create, read, update, delete*) Representa las operaciones básicas que debe tener la implementación de una base de datos, las cuales son: crear, leer, actualizar y eliminar la información de la misma.
- **Búsqueda Intuitiva:** Característica de un método de búsqueda el cual brinda posibles resultados antes de que los parametrizables sean especificados completamente.
- **Complejidad algorítmica:** Concepto matemático que describe la eficiencia en recursos informáticos a la hora de realizar una secuencia de operaciones.
- **Árbol binario de búsqueda auto balanceado:** En informática, es una estructura para gestionar datos la cual nos ofrece una complejidad garantizada $O(\log n)$ en sus principales operaciones: búsqueda, inserción y eliminación.
- **Trie:** Es un concepto teórico de árbol que brinda un almacenamiento eficiente de nodos conectados a través de un conjunto de claves las cuales están compuestas por una secuencia de caracteres pertenecientes a un alfabeto dado.

FASE 3: BÚSQUEDA DE SOLUCIONES CREATIVAS

Para la generación de ideas se realizó una revisión secuencial a conceptos previamente adquiridos relacionados con las necesidades del problema anteriormente expuesto, tales como: eficiencia algorítmica, estructuras de datos, generación automatizada de datos, etc. Partiendo de lo anterior, las posibles soluciones encontradas se exponen a continuación:

- **Generación automatizada de personas:**
 - a. Elaborar módulo de software que genere información aleatoria a partir de un data set diseño propio.
 - b. Implementación de data sets previamente creados y ofrecidos por plataformas digitales tales como: data.world, [indexmundi](https://indexmundi.com), [kaggle](https://kaggle.com), [thispersondoesnotexist](https://thispersondoesnotexist.com).
- **Modelamiento de una estructura de datos eficiente:**
 - a. Implementar una estructura Hash Table dinámica para almacenar los datos.
 - b. Implementar una estructura de árbol de búsqueda binario para el almacenamiento de los datos.
 - c. Implementar una estructura de árbol de búsqueda binario auto balanceado para el almacenamiento de los datos.
 - d. Implementar una estructura de árbol de búsqueda rojinegro para el almacenamiento de los datos.
 - e. Implementar una estructura de grafos para el almacenamiento de los datos.
 - f. Implementar estructuras de datos basadas en el concepto de trie
- **Búsqueda intuitiva:**
 - a. Implementación de estructuras teóricas como grafos, que brindan un recorrido parcial dada una secuencia de símbolos.
 - b. Implementación de estructuras teóricas como árbol binario de búsqueda.
 - c. Solicitar al usuario el dato a buscar y luego con este dato buscarlo en la base de datos.
- **Eliminación de datos:**
 - a. Hacer uso de la solución implementada para búsqueda intuitiva con el fin de obtener el dato eliminado correctamente.

- b. Solicitar al usuario el dato a eliminar y luego removerlo de la base de datos.
- **Actualizar datos**
 - a. Hacer uso de la solución implementada para búsqueda intuitiva con el fin de obtener el dato y actualizar su información.
 - b. Solicitar al usuario el dato modificado y éste reemplazarlo en la base de datos.
 - c. Hacer uso de la solución implementada para la búsqueda intuitiva con el fin de obtener el dato y actualizar su información con información nueva generada de manera automatizada.

FASE 4: TRANSICIÓN DE LA FORMULACIÓN DE IDEAS A LOS DISEÑOS PRELIMINARES

En esta fase hemos descartado las peores alternativas que no brindan una solución adecuada a los requerimientos, tales como: el modelamiento de estructura de datos de tipo Hash Table para el almacenamiento de la información, posteriormente se descarta la idea de solicitar al usuario el elemento a buscar previamente y en este mismo orden de ideas se hace lo mismo para la eliminación y la actualización de datos.

- **Generación automatizada de personas:**

Analizando a profundidad lo requerido, todas las ideas propuestas brindan gran potencial para generar automáticamente personas, por lo tanto, ninguna se descarta aún.

- **Modelamiento de una estructura de datos eficiente:**

Dándole una mirada más profunda a la estructura Hash Table no cumple con lo requerido para esta solución dado que no es lo suficientemente eficiente para implementarlo como base de una búsqueda predictiva, por lo tanto, se descarta esta idea para el modelamiento de una estructura de datos eficientes.

- **Búsqueda intuitiva:**

Dado que no se ha recolectado la suficiente información sobre grafos ni árboles AVL o rojinegros dichas ideas aún no se descartan. En el caso de la idea de solicitar previamente al usuario el dato a buscar se descarta dado que no es eficiente y no brinda una lógica racional.

- **Eliminación de datos:**

La idea de solicitar previamente al usuario el dato a eliminar se descarta dado que no es eficiente y no brinda una lógica racional.

- **Actualizar datos:**

La idea de actualizar la información con información generada de manera automatizada se descarta dado que no es eficiente y no brinda una lógica racional.

FASE 5: EVALUACIÓN Y SELECCIÓN DE LA MEJOR SOLUCIÓN

- **Criterios generación automatizada de personas**
 - a. Permita hacer uso de data sets previamente definidos
 - b. Permita incluir parámetros deliberados para la generación de datos
- **Criterios modelamiento de una estructura eficiente**
 - a. Permita almacenar los datos de manera eficiente
 - b. Permita almacenar una gran cantidad de datos

- c. Permita un acoplamiento adecuado a nuestra propuesta de implementación para la búsqueda de datos intuitiva
- **Búsqueda intuitiva**
 - a. Permita implementar una búsqueda teórica mediante claves de secuencia de caracteres pertenecientes a un alfabeto dado (tier).
 - b. Que ofrezca una complejidad espaciotemporal eficiente.
- **Eliminación de datos**
 - a. Llevar acabo la eliminación simple de un dato dado
- **Modificación de datos**
 - a. Llevar acabo la modificación simple de un dato dado

Selección alternativas generación automatizada de personas

	Criterio a	Criterio b	Total
Alternativa a	0	5	5
Alternativa b	5	0	5

No se descarta ninguna opción dado que ambas tienen el mismo total, por lo que se implementarán las dos opciones.

Selección alternativas modelamiento de una estructura de datos eficiente

	Criterio a	Criterio b	Criterio c	Total
Alternativa b	3	5	5	13
Alternativa c	5	5	5	15
Alternativa d	5	5	5	15
Alternativa e	5	5	5	15
Alternativa f	5	5	5	15

La alternativa b se descarta parcialmente, pero se implementarán algunas de sus definiciones teóricas para la construcción de las demás alternativas.

Selección alternativas búsqueda intuitiva

	Criterio a	Criterio b	Total
Alternativa a	5	5	10
Alternativa b	0	5	5

La alternativa b se descarta parcialmente, pero se implementarán algunas de sus definiciones teóricas para la construcción de las demás alternativas.

Selección alternativas eliminación de datos

	Criterio a	Total
Alternativa a	5	5

Dado que la única alternativa es a elegimos esta idea

Selección alternativas actualización de datos

	Criterio a	Total
Alternativa a	5	5
Alternativa b	3	3

Se descarta la idea b por ser la de menor puntaje total