# Observaciones laboratorio 6

# Teniendo en cuenta cada uno de los requerimientos ¿Cuántos índices implementaría en el Reto? y ¿Por qué?

* Para el requerimiento 1, habría que implementar un índice por año de nacimiento para los artistas. Ya con este índice se puede escoger las listas de autores que tienen por llave años dentro del rango que se quiere ordenar. De la misma forma, se tendrá que hacer un índice con fechas de adquisición para el requerimiento 2 que funcionará de la misma manera para seleccionar las obras adquiridas dentro de un rango de fechas. Para el requerimiento 3, yo implementaría en el catálogo un índice que tenga como llaves los nombres de los artistas y como valores sus diccionarios para disminuir la complejidad de buscar el id de un artista, además de un índice que tenga como llaves los id´s de los artistas de las obras y como valores todas las obras que tengan ese artista entre sus artistas, para así encontrar la lista de obras del artista que se busca, y con esta hacer un mapa auxiliar con un índice que use como llaves los medios usados por el artista. Para el requerimiento 4, usaremos el mismo índice que ordena las obras por sus artistas. Así, para cada llave en el índice de artistas, se buscará la nacionalidad de ese artista y se agregarán todas las obras de la lista del id de ese artista a la nacionalidad correspondiente en un nuevo índice que tiene como llaves las nacionalidades. Para el requerimiento 5, vamos a hacer un índice que tenga como llaves los departamentos del museo y como valores todas las obras pertenecientes a esos departamentos. Por último, para el requerimiento 6 usaremos el mismo índice que tiene como llaves los artistas y para cada artista haremos un mapa auxiliar que tiene como llaves los diferentes medios.
* Así tendríamos que habría que implementar 6 índices en el catálogo del reto:

\*Por fecha de nacimiento de los artistas

\*Por fecha de adquisición de las obras

\*Por id de los artistas con sus diccionarios como valores

\*Por nombre del artista con una lista con sus obras como valor

\*Por nacionalidades que tiene como valor la lista de obras

\*Por departamentos del museo

* Además, para los requerimientos 3 y 6 se crearán mapas fuera del catálogo que tienen como llaves los medios y como valores listas de obras, pero solo de un artista específico por mapa.
* Usar todos estos índices implementarían una solución con tiempos de respuesta muy bajos, pero que ocupa mucha memoria implementando todos los índices en el catálogo. Para una solución equilibrada entre gastos de espacio y tiempo, tendríamos que quitar algunos índices que nos dan una mejora en el tiempo.

## Según los índices propuestos ¿en qué caso usaría Linear Probing o Separate Chaining en estos índices? y ¿Por qué?

* Habría que usar separate chaining en todos los índices que engloben varios elementos y los agrupen en categorías, pues es más cómodo tener todos los elementos que colisionan por tener la misma llave agrupados en una sola lista que hace más fácil buscarlos por la característica de su llave. Por otro lado, usaremos linear probing en los índices que no buscan agrupar elementos, sino que tienen un solo elemento en valor y no se deberían formar listas muy grandes.
* En este caso usaríamos Separate Chaining en:

\*Índice por nombre del artista con una lista con sus obras como valor

\*Índice por nacionalidades que tiene como valor la lista de obras

\*Índice por departamentos del museo

\*Índices auxiliares por medio para las obras de un artista específico

* Los índices por fecha de nacimiento del artista y fecha de adquisición de la obra parece que no tuviesen muchas colisiones con pocos datos, pero analizando los datos de las fechas y mediante se va aumentando el tamaño de los datos, se van generando más colisiones de obras adquiridas en la misma fecha y, sobretodo, de artistas nacidos en el mismo año, por lo que también usaremos Separate Chaining.
* Usaríamos Linear Probing en:

\*Índice por id de los artistas con sus diccionarios como valores

## Dado el número de elementos de los archivos MoMA, ¿Cuál sería el factor de carga para estos índices según su mecanismo de colisión?

* Los factores de carga para los índices (suponiendo que las colisiones sean uniformes) serían:

\*Índice por nombre del artista con una lista con sus obras como valor: el número de obras de un artista sobre el número de artistas.

\*Índice por nacionalidades que tiene como valor la lista de obras: el número de obras que tienen algún artista de una nacionalidad específica sobre el número de nacionalidades en el catálogo.

\*Índice por departamentos del museo: el número de obras para un departamento sobre el número total de departamentos.

\*Índices auxiliares por medio para las obras de un artista específico: el número de obras de un autor específico hechas por un medio específico sobre el número total de medios que usó el autor en sus obras.

\*Índice por año de nacimiento de autores: el número de autores nacidos en un mismo año sobre el número total de años en los que nació algún autor del catálogo.

\*Índice por fecha de adquisición de las obras: el número de obras adquiridas en una misma fecha sobre el número total de fechas en las que se adquirió una obra del catálogo.

\*Índice por id de los artistas con sus diccionarios como valores: en este caso, el factor sería el número de artistas sobre la capacidad de llaves del mapa. Sería el factor de carga más bajo.