

## Trabajo Práctico Nº 1

**Tema:** Introducción

**Fecha Inicio:** 26/03/2024    **Fecha de Entrega:** 09/04/2024

### Actividades:

- 1) Utilizando como recurso el motor de búsqueda Google trate de satisfacer las siguientes necesidades de información. Indique en cada caso la cantidad de consultas que tuvo que realizar para llegar al objetivo (si llegó), cuántas referencias de la respuesta revisó en cada caso y en cuál encontró lo buscado (si es necesario, puede escribir las consultas en inglés). Suponga que usted no sabe la respuesta a ninguna de las preguntas.
  - a) ¿Cuál fue el año del primer juego olímpico?
  - b) ¿Qué motivó la caída del muro de Berlín? En qué año sucedió?
  - c) ¿Cuál fue el ranking de los millonarios el año pasado?
- 2) Ahora realice las mismas consultas del punto 1) pero empleando ChatGPT y Gemini
- 3) Cree una aplicación que “indexe” los siguientes 5 documentos. Debe formar una colección con todas las palabras, asegúrese de eliminar las [stop-words](#). La aplicación debe leer los documentos .PDF previamente descargados a una ubicación del disco rígido y armar un fichero invertido con frecuencia y por separado un fichero invertido posicional. Con la última representación, genere 4 consultas distintas y muestre los resultados.
  - a. <https://sedici.unlp.edu.ar/handle/10915/155979>
  - b. <https://sedici.unlp.edu.ar/handle/10915/162331>
  - c. <https://sedici.unlp.edu.ar/handle/10915/162045>
  - d. <https://sedici.unlp.edu.ar/handle/10915/152364>
  - e. <https://sedici.unlp.edu.ar/handle/10915/158824>
- 4) El siguiente es un resumen que se encuentra construido con un índice invertido, se indica el tamaño del mismo y cada palabra posee entre corchetes la posición que ocupa dentro del texto. Cree una aplicación que permita recibir estos dos parámetros, el tamaño del índice y el índice invertido, y devuelva el resumen completo. Cree las estructuras de datos que considere necesario tanto para la re-construcción del texto como para el pasaje de parámetros.

**Tamaño del Índice:** 184

**Índice Invertido:**

```
{ "COVID-  
19, ": [0], "caused": [1], "by": [2], "the": [3, 16, 26, 31, 36, 54, 62, 70, 78, 85, 97, 110, 117  
, 125, 146, 162, 173, 178], "coronavirus": [4], "SARS-  
CoV2, ": [5], "continues": [6], "to": [7, 58, 73, 77, 129], "pose": [8], "a": [9, 170], "glob  
al": [10, 37], "threat": [11], "The": [13], "consequences": [14], "from": [15, 116], "pan  
demic": [17], "are": [18], "devastating, ": [19], "not": [20, 141], "only": [21], "for": [2  
2, 35, 124, 172], "human": [23], "health": [24, 28], "and": [25, 43, 82, 103, 132], "nation  
al": [27], "systems": [29], "throughout": [30], "world, ": [32], "but": [33], "also": [34  
, 107], "economy": [38], "Bats, ": [40], "pangolins, ": [41], "muskrats": [42], "other": [4  
4, 136], "wild": [45], "animals": [46, 128, 137], "have": [47, 74], "been": [48, 122], "im  
plicated": [49], "as": [50, 87], "possible": [51], "hosts": [52], "of": [53, 65, 84, 93, 10
```

0,112,127,148,164,175], "virus": [55], "According": [57], "recent": [59], "research": [60], "data,": [61], "increased": [63], "concentration": [64], "particulate": [66], "air": [67,94], "matter": [68], "in": [69,145,152,177], "atmosphere": [71], "appears": [72,108], "contributed": [75], "significantly": [76], "spread": [79], "(airborne": [80], "transmission)": [81], "aggressiveness": [83], "disease,": [86], "large": [88], "cities": [89], "with": [90,135], "high": [91], "levels": [92], "pollution": [95], "showed": [96], "highest": [98], "numbers": [99], "SARS-CoV2": [101], "cases": [102], "deaths": [104], "It": [106,157], "that": [109,138,161], "collapse": [111], "natural": [113], "ecosystems": [114], "resulting": [115,144], "ever-evolving": [118], "climate": [119], "change": [120], "has": [121,181], "crucial": [123], "migration": [126], "new": [130,149], "areas": [131], "their": [133], "interaction": [134], "they": [139], "would": [140], "normally": [142], "encounter,": [143], "emergence": [147], "pathogens,": [150], "which": [151,167], "turn": [153], "infected": [154], "humans": [155], "is": [158], "now": [159], "clear": [160], "era": [163], "climate": [165], "medicine": [166], "will": [168], "be": [169], "cornerstone": [171], "practice": [174], "medicine": [176], "21st": [179], "century,": [180], "already": [182], "begun": [183]}