



# Cómputo Paralelo

y en la Nube

*Proyecto*

*Clausura*

*Bolsa de Palabras con*  
**MPI** para el Análisis de  
texto

# Escenario con 6 Libros para ejecución con 6 Procesos

- `shakespeare_the_merchant_of_venice`
- `shakespeare_romeo_juliet`
- `shakespeare_hamlet`
- `shakespeare_the_merchant_of_venice`
- `shakespeare_romeo_juliet`
- `shakespeare_hamlet`

	...	yourn	youth	youthful	youthfulness	zeal	zealous	zenith	zip	zone	zounds
0	...	0	8	1	0	1	0	0	0	0	0
1	...	0	6	3	0	0	0	0	1	0	2
2	...	0	14	0	0	0	0	0	2	1	0
3	...	0	1	0	0	1	0	0	1	0	0
4	...	0	9	6	0	0	2	1	0	0	0
5	...	1	10	3	1	0	2	0	1	0	0







# Requerimientos

## Entrada

- Listado de nombres de los archivos donde se encuentran las palabras a contar (los archivos se pueden encontrar en el mismo lugar que el ejecutable)
- Nombre del archivo en donde se encuentra el vocabulario y su tamaño.
- Número de procesos a utilizar que debe ser igual al número de archivos a introducir

## Salida

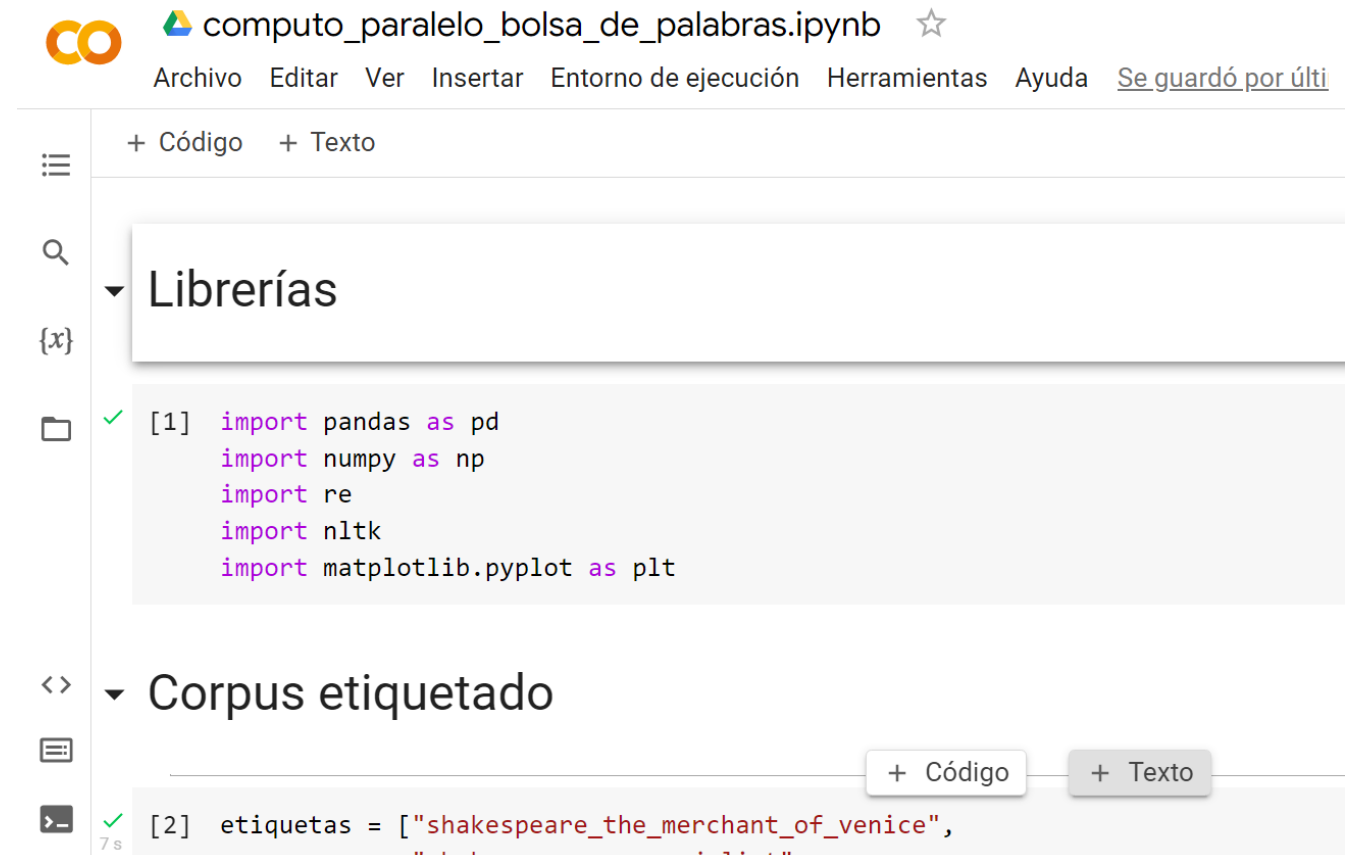
- Archivo con matriz de Bolsa de Palabras en formato csv.

# Requerimientos

- Para simplificar el problema, el código debe poder trabajar exactamente para 6 libros, pero pueden ser cualquier libro.
- Para simplificar el problema, se puede asumir el vocabulario y su tamaño (se puede extraer de la libreta de python).
- Implementar versión paralela con MPI y Versión Serial.
- Comparar el tiempo de ejecución de la version paralela con la versión serial.
- Obtener gráfica de Speed Ups

# Libreta de Google Colab

- [https://colab.research.google.com/drive/1v4VvVDYpdeZjtnDcUzKeSVjDk5W6\\_YNf?usp=sharing](https://colab.research.google.com/drive/1v4VvVDYpdeZjtnDcUzKeSVjDk5W6_YNf?usp=sharing)



# Criterios de evaluación

- Peso total del proyecto: 2.5 puntos de su calificación final.
- Ejecución del proyecto con todos los requerimientos indicados en su descripción (2.3 puntos)
- Código documentado correctamente (0.2 puntos)
- Posibles fecha de entrega:
  - Martes 29 de Noviembre de 2022 en clase.  
Equipo de 3 Personas
- **NOTA 1:** Si el código paralelo no es más eficiente que el código en automático aplica un descuento de **1 punto.**
- **NOTA 1:** 30% menos por cada día natural de retraso.
- **NOTA 2:** Si se entrega después de la hora de entrega, en automático aplica un día menos.

# Hola Mundo Diccionarios

```
#include <map>

#include <string>

#include <iostream>

using namespace std;

int main (int argc, char *argv[]) {

    map<string, int> my_dictionary;

    my_dictionary.insert({"computo", 100});

    my_dictionary.insert({"paralelo", 200});

    my_dictionary.insert({"curso", 300});

    //my_dictionary.at("paralelo") = 3;

    my_dictionary["paralelo"] = 4;

    cout << my_dictionary.at("paralelo") << "\n";

    cout << my_dictionary.at("computo") << "\n";

    cout << my_dictionary["paralelo"] << "\n";

    cout << my_dictionary.count("inexistente") << "\n";

    cout << my_dictionary.count("paralelo") << "\n";

    my_dictionary.erase("curso");

    for (auto iterator = my_dictionary.begin(); iterator != my_dictionary.end(); ++iterator) {

        cout << iterator->first << " " << iterator->second << "\n";

    }

    return 0;

}
```