

Presentación del Proyecto MoogLe

Pablo Gómez Vidal

MATCOM
Universidad de la Habana

Julio, 2023

¿Qué es el Moogle?

Moogle! es una aplicación "totalmente original" cuyo propósito es buscar inteligentemente un texto en un conjunto de documentos. Es una aplicación web, desarrollada con tecnología .NET Core 6.0, específicamente usando Blazor como framework web para la interfaz gráfica, y en el lenguaje Csharp.

¿Qué puede hacer?

- Busca de manera local archivos en formato.txt
- Con una base de datos de 100 mb realiza las búsquedas en menos de 1 segundo
- Cuenta con un Snippet que enseña un trozo del documento donde aparece la palabra de la búsqueda
- Cuenta con una sugerencia a la palabra mas parecida a la de la búsqueda en caso de que esa no sea encontrada
- Cuenta con hipervínculo hacia el documento que aparece en la búsqueda

¿Cómo funciona?

Funciona de manera simple, primero se leen las palabras de todos mis documentos y se tokenizan, luego se calcula el valor de importancia de cada palabra y se guarda para no tener que repetir el proceso, luego al introducir el query en la búsqueda se le hace el mismo proceso de tokenizado y se busca en cada documento el valor de esa palabra (query) y se escogen los 5 documentos con mayor score mediante una fórmula llamada similitud de coseno, y se imprime el resultado de búsqueda