**Tecnológico de Costa Rica**

**Escuela de Ingeniería en Computación**

**IC-8020 Recuperación de Información Textual**



**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Tarea programa: Índice invertido**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Profesor:**

Jose Enrique Araya Monge

**Estudiantes:**

Mario Naranjo Leiva, 2013034348

Marcelo Guzmán Rivera, 201135367

II Semestre, 2016

**Contenido**

[Introducción 1](#_Toc462865993)

[Estado del proyecto 1](#_Toc462865994)

[Pruebas 2](#_Toc462865995)

# Introducción

# Estado del proyecto

|  |  |  |
| --- | --- | --- |
| **Etapa** | **% de complet.** | **Comentario o aclaración** |
| **Indexa toda la colección** |  |  |
| **Transformaciones implementadas** |  |  |
| Indexa palabras (conversión a minúsculas, eliminación de acentos, preservaión de eñe) |  |  |
| Usa lista de stopwords indicada |  |  |
| Indexa números |  |  |
| Indexa frases pedidas |  |  |
| **Estructura archivo invertido** |  |  |
| Archivo documentos (dar # total docs.) |  |  |
| Diccionario (dar # total términos) |  |  |
| Postings (tamaño fijo, dar # total de entradas) |  |  |
| **Implementa consultas vectoriales** |  |  |
| Permite términos simples |  |  |
| Permite ajustes +/- en consultas |  |  |
| Permite prefijos (\*) |  |  |
| Permite frases (“) |  |  |
| Crea archivo escalafón correctamente |  |  |
| **Implementa consultas avanzadas** |  |  |
| Procesa una cláusula: <entidad> <caracter> <valor> |  |  |
| Procesa más de una cláusula (AND) |  |  |
| Crea archivo escalafón correctamente |  |  |
| **Documento salida** |  |  |
| Genera archivo HTML de salida |  |  |
| Incluye todos los datos pedidos |  |  |
| **Interfaz** Indicar estado de la implementación de los siguientes aspectos: |  |  |
| Indexar una colección |  |  |
| Realizar consulta vectorial |  |  |
| Realizar consulta estructurada |  |  |
| Mostrar resultados |  |  |

# Pruebas

| **Pruebas** | **Comentarios sobre resultados obtenidos vs resultados esperados** |
| --- | --- |
| Prueba 1 |  |
| Prueba 2 |  |
| Prueba 3 |  |
| Prueba 4 |  |
| Prueba 5 |  |
| Prueba 6 |  |
| Prueba 7 |  |
| Prueba 8 |  |
| Prueba 9 |  |
| Prueba 10 |  |