

UNIVERSIDAD DE COSTA RICA
ESCUELA DE CIENCIAS DE LA COMPUTACIÓN E INFORMÁTICA
CI-2414 RECUPERACIÓN DE INFORMACIÓN
PROF. KRYSZIA DAVIANA RAMÍREZ BENAVIDES

INVESTIGACIÓN

OBJETIVO

- Introducir al estudiante a la investigación en el área de Recuperación de Información mediante el desarrollo de una investigación y una aplicación supervisada en un tema específico.

FECHAS DE ENTREGA

- Entrega de la Presentación ⇒ Una semana antes de la exposición.
- Exposición de la Aplicación ⇒ Del 14 de septiembre al 16 de noviembre.

EVALUACIÓN

Exposición del Tema y la Aplicación	15%
Evaluaciones de las Exposiciones	5%
Total	20%

NOTAS IMPORTANTES

- La investigación consiste en la presentación al grupo, en grupo de dos personas como máximo, de una aplicación sobre temas vistos por el profesor en clases o sobre temas que deben investigar.
- Se formarán los grupos y se asignarán los temas el primer día de clases.
- La presentación debe ser entregada una semana antes de la fecha de exposición, para que sea revisada por el profesor.
- La aplicación debe ser hecha por el grupo de estudiantes.
- Cada estudiante llenará una hoja con comentarios sobre las exposiciones de cada grupo, estos comentarios serán evaluados y pasarán a formar parte de su nota individual; ya que indican el interés, aprovechamiento y asimilación de los temas expuestos.

ASPECTOS METODOLÓGICOS

Partiremos del interés personal de los estudiantes para asignar una serie de temas que serán desarrollados por cada grupo de estudiantes. La investigación, la presentación y la aplicación se harán en grupos de dos personas. Cada grupo desarrollará su tema y una aplicación del mismo, y culminará con una presentación en forma de exposición en clase.

TEMAS DE INVESTIGACIÓN

- Otros SRI: Yahoo y Bing (Kumo), y comparación con Google (teoría)
- Aplicación – *Crawler*.
- Aplicación – Analizador léxico.
- Aplicación - Algoritmos de similaridad y distancia:
 - Distancia de Hamming.
 - Distancia de Levenstein.
 - Bigramas.
- Aplicación – Lematizador:
 - Porter.
 - Variedad de sucesores.
 - Bigramas.
- Aplicación – Clustering:
 - K-Means.
 - HAC.
- Aplicación – Interface de visualización:
 - Tree Map.
- Aplicación – Compresor:
 - Huffman orientado a bytes.
- Aplicación – Filtrador de Documentos (teoría y ejemplos).
- Aplicación – Sumarización (teoría y ejemplos).
 - Algoritmo Extractor.
- Optimización en Buscadores (SEO: Search Engine Optimization) (teoría y ejemplos).
- RI Multimedia (sonidos, imágenes, videos) (teoría y ejemplos).