



Ministerio de Cultura y Educación  
 Universidad Nacional de San Luis  
 Facultad de Ciencias Físico Matemáticas y Naturales  
 Departamento: Informatica  
 Area: Area I: Datos

(Programa del año 2022)

### I - Oferta Académica

Materia	Carrera	Plan	Año	Período
() RECUPERACIÓN DE LA INFORMACIÓN	ING. EN COMPUT.	28/12	2022	2° cuatrimestre
		026/1		
() RECUPERACIÓN DE LA INFORMACIÓN	ING. INFORM.	2-	2022	2° cuatrimestre
		08/15		
() RECUPERACIÓN DE LA INFORMACIÓN	LIC.CS.COMP.	18/11	2022	2° cuatrimestre
() RECUPERACIÓN DE LA INFORMACIÓN	LIC.CS.COMP.	32/12	2022	2° cuatrimestre

### II - Equipo Docente

Docente	Función	Cargo	Dedicación
HERRERA, NORMA EDITH	Prof. Responsable	P.Asoc Exc	40 Hs
LUDUEÑA, VERONICA DEL ROSARIO	Prof. Colaborador	P.Adj Exc	40 Hs
REYES, NORA SUSANA	Prof. Colaborador	P.Asoc Exc	40 Hs

### III - Características del Curso

Credito Horario Semanal				
Teórico/Práctico	Teóricas	Prácticas de Aula	Práct. de lab/ camp/ Resid/ PIP, etc.	Total
3 Hs	Hs	Hs	2 Hs	5 Hs

Tipificación	Periodo
B - Teoria con prácticas de aula y laboratorio	2° Cuatrimestre

Duración			
Desde	Hasta	Cantidad de Semanas	Cantidad de Horas
08/08/2022	18/11/2022	15	75

### IV - Fundamentación

A partir de la expansión y consolidación de Internet se ha puesto a disposición de las personas una importante cantidad de información de todo tipo. En la actualidad, la cantidad de información no estructurada que se genera y distribuye supera ampliamente las posibilidades de los usuarios de realizar un uso eficiente.

Para poder aprovechar al máximo este potencial de información, se necesitan modelos, algoritmos y técnicas que permitan su gestión eficaz y eficiente. En esta materia se abordan los fundamentos teóricos de la temática de Recuperación de la Información junto con el estudio de las herramientas que permiten la construcción de un sistema de estas características.

### V - Objetivos / Resultados de Aprendizaje

Al completar la asignatura, se espera que los estudiantes sean capaces de:

- Comprender los principales desafíos de recuperación de información y sus diferencias con otras disciplinas relacionadas
- Conocer conceptos y modelos fundamentales de recuperación de información.
- Comprender los fundamentos teóricos sobre los modelos clásicos de recuperación de información
- Comprender los métodos de indexación usados en este ámbito.

- Comprender los alcances de la disciplina, junto con criterios que les permitan determinar sus ámbitos de aplicación.

## VI - Contenidos

### Unidad 1: Introducción

Introducción al problema de recuperación de información. Diferencias con el concepto de recuperación de datos. Arquitectura de un sistema de recuperación de información.

### Unidad 2: Modelos

Modelos de recuperación de la información. Una taxonomía de modelos. Modelos clásicos: booleano, vectorial y probabilístico. Modelos alternativos.

### Unidad 3: Preprocesamiento e Indexación

Análisis y preprocesamiento de textos. Análisis estadístico de las propiedades del texto. Procesamiento de documentos y términos. Eliminación de stopwords y lematización. Indexación: listas invertidas y archivos signatures

### Unidad 4: Consultas.

Consultas basadas en palabras claves. Pattern matching. Expansión de una consulta y retroalimentación. Expansión basada en clustering y expansión basada en thesaurus. Evaluación de la recuperación: exhaustividad (Recall) y precisión (Precision). Diagramas de Exhaustividad/Precisión.

### Unidad 5: Recuperación de Información en la Web

Características de la recuperación de información en la Web. Arquitectura de los motores de búsquedas. Algoritmos de ranking. Indexación y crawling.

## VII - Plan de Trabajos Prácticos

Los trabajos consisten en la realización de las siguientes actividades: resolución de ejercicios, presentaciones de temas asignados por los docentes, evaluación crítica del software disponible, elaboración de informes sobre los temas estudiados, búsqueda de material disponible en Internet, estudio de un software específico para recuperación de información.

## VIII - Regimen de Aprobación

Para regularizar el alumno deberá:

- 1) Asistir al 80% de las clases.
- 2) Entregar y aprobar los trabajos prácticos asignados por los docentes.
- 3) Aprobar un examen (oral o escrito), o alguna de sus dos recuperaciones.

Para Promocionar deberá regularizar, aprobando el examen con una nota de 7 puntos o más, y deberá realizar un trabajo integrador sobre los temas estudiados. La nota final de promoción será el promedio de las notas obtenidas en el examen y en el trabajo integrador.

Modalidad de examen final: El examen final podrá ser oral y/o escrito.

Examen Libre: No se admiten alumnos libres por la modalidad de la asignatura.

## IX - Bibliografía Básica

- [1] R. Baeza-Yates, B. Ribeiro-Neto. Modern Information Retrieval. 2nd Ed. Addison-Wesley, 2011.
- [2] S. Büttcher, C.L.A. Clarke, G.V. Cormack. Information Retrieval: Implementing and Evaluating Search Engines. The MIT Press, 2016.
- [3] M. McCandless, E. Hatcher and O. Gospodnetic. Lucene in Action (2d Edition) Manning 2011.
- [4] Stefano Cerri r Alessandro Bozzon, Marco Brambilla , Emanuele Della Valle, Piero Fraternali, Silvia Quarteroni. Web Information Retrieval. Ed. Springer 2013.

## X - Bibliografía Complementaria

## **XI - Resumen de Objetivos**

- Comprender los principales desafíos de recuperación de información y sus diferencias con otras disciplinas relacionadas
- Conocer conceptos y modelos fundamentales de recuperación de información.
- Comprender los fundamentos teóricos sobre los modelos clásicos de recuperación de información
- Comprender los métodos de indexación usados en este ámbito.
- Comprender los alcances de la disciplina, junto con criterios que les permitan determinar sus ámbitos de aplicación.

## **XII - Resumen del Programa**

Unidad 1: Introducción

Unidad 2: Modelos

Unidad 3: Preprocesamiento e Indexación

Unidad 4: Consultas. Evaluación: precisión y recall.

Unidad 5: Recuperación de Información en la Web

## **XIII - Imprevistos**

## **XIV - Otros**