

# PERTEMUAN 1



# Buku Bacaan

- C. Manning, P. Raghavan, and H. Schutze. Introduction to Information Retrieval. Cambridge University Press, 2008.
- Jurafsky, Martin. Speech and Language Processing. Prentice Hall, 1999.

# Materi Pembelajaran 1

- Deskripsi Singkat konsep information retrieval.
- Peran Mata Kuliah Topik dalam Information Retrieval dalam kerangka Kurikulum Program
- Tujuan Instruksional Umum dan Khusus
- Aturan Perkuliahan

# Materi Pembelajaran 2

## **Pengertian Information Retrieval**

- Pengantar konsep information retrieval
- Perbedaan antara information retrieval, question answering, information extraction, text summarization dan text classification
- Arsitektur umum pada information retrieval
- Isu-isu penting dalam information retrieval

# Materi Pembelajaran 3

## **Arsitektur Umum Sistem Information Retrieval**

- Arsitektur sistem information retrieval secara umum
- Sub sistem crawling pada IR system
- Sub sistem indexing pada IR system
- Sub sistem pencarian relevant document pada IR

# Materi Pembelajaran 4

## **Lexical processing**

- tokenisasi
- normalisasi
- stemming
- pembobotan kata
- stop word elimination
- word sense disambiguation
- thesaurus

# Materi Pembelajaran 5

## **Vector Space Model**

- Vector space model
- Index construction
- MapReduce

# Materi Pembelajaran 6

## **Tolerant Retrieval**

- Dictionary data structure
- Tolerant Retrieval



# Materi Pembelajaran 7

## **Evaluasi sistem IR**

- Precision
- Recall
- Accuracy
- F-Measure
- Mean average precision (MAP)
- R-precision
- Variance
- Kappa Measure
- Result presentation

# Materi Pembelajaran 8

## **Query Expansion**

- Related term taken from relevant documents.
- Using Thesaurus.

# Materi Pembelajaran 9

## **Lucene tutorial**

- Setiap fungsi pada lucene

# Materi Pembelajaran 10

## **Document clustering**

- Penggunaan clustering pada IR system
- Jenis clustering: hard vs soft
- K-mean clustering
- Agglomerative clustering
- Fuzzy clustering

# Materi Pembelajaran 11

## **Probabilistic Model untuk Information Retrieval**

- Probability ranking principle
- Binary independence model
- Tree structured dependencies between term
- Okapi BM 25
- Bayesian network approach to IR

# Materi Pembelajaran 12

## **Latent Semantic Indexing**

- Teknik latent semantic indexing
- Penggunaannya pada IR system

# Materi Pembelajaran 13

## **Named Entity Recognition pada IR System**

- Teknik NER
- Penggunaan NE pada IR System

# Materi Pembelajaran 14

Case study



**TERIMA KASIH**