

Programma di Sistemi per il recupero delle informazioni

Gabriele Pozzani

Corso di Laurea Magistrale in Editoria e Giornalismo, A.A. 2011/2012
Università degli Studi di Verona, Italia

Argomento 1. Introduzione ai sistemi informativi e ai sistemi di recupero delle informazioni

Fonte: dispense pubblicate sulla pagina Web del corso

- introduzione ai sistemi informativi
- informazioni e dati
- tipi di sistemi informativi
- information retrieval
- architettura di un sistema di IR
- criteri di efficacia

Argomento 2. Tipi di interrogazioni

Fonte: capitolo 3 di ISR¹

- booleane
- vettoriali
- booleane estese
- fuzzy
- probabilistiche
- in linguaggio naturale

Argomento 3. Il processo di matching

Fonte: capitolo 4 di ISR¹

- misure di rilevanza e similarità
- matching booleano
- matching vettoriale
- matching fuzzy
- matching probabilistico
- matching basato sulla prossimità

Argomento 4. Analisi dei testi

Fonte: capitolo 5 di ISR¹

- rappresentazione matriciale
- indicizzazione
- legge di Zipf
- stop list
- thesauri (*Fonte: dispense pubblicate sulla pagina Web del corso*)

Argomento 5. Prestazioni dei sistemi di recupero

Fonte: capitolo 8 di ISR¹

- precisione e richiamo
- fallout e generalità
- misure orientate all'utente
- lunghezza di ricerca attesa

Argomento 6. XML per l'editoria elettronica

Fonte: slides pubblicate sulla pagina Web del corso

- introduzione ad XML
- documenti orientati alla narrazione
 - XML come formato documentale
 - struttura dei documenti XML orientati alla narrazione
 - TEI (Text Encoding Initiative)
 - DocBook
- documenti orientati ai dati
 - XML come formato per i dati
 - applicazioni XML per la programmazione

¹R. Korfage, *Information Storage and Retrieval*, Publisher John Wiley & Sons

Capitoli e sezioni di ISR¹ che sono parte del programma

Argomento 2. Tipi di interrogazioni: capitolo 3 sezioni 3.1, 3.2, 3.3 (senza formula a pag. 65), 3.4 (fino a "... in S to be considered"), 3.5, 3.6, 3.7;

Argomento 3. Il processo di matching: capitolo 4 sezioni 4.1, 4.2, 4.3 (fino a pag. 83 riga 8), 4.4, 4.6, 4.7, 4.8;

Argomento 4. Analisi dei testi: capitolo 5 sezioni 5.1, 5.2, 5.3 (fino a pag. 113 riga 10), 5.5 (fino a pag. 116 riga 18), 5.8, 5.9, 5.11 (più dispense su thesauri pubblicate sulla pagina web del corso);

Argomento 5. Prestazioni dei sistemi di recupero: capitolo 8 sezioni 8.2, 8.3, 8.6 (fino all'esempio 8.2 compreso).