

Sistemi per il recupero delle informazioni

Laurea Magistrale in Editoria e Giornalismo

Prova scritta del 16 febbraio 2015

Cognome e nome: _____ Matricola: _____

Domanda 1	Domanda 2	Domanda 3	Domanda 4	Domanda 5	Domanda 6	Domanda 7	Totale
/5	/5	/5	/5	/5	/5	/3	/33

Istruzioni:

- ◆ È vietato portare all'esame libri, eserciziari, appunti e dispense. Chiunque venga trovato in possesso di documentazione relativa al corso, in formato analogico e/o digitale, – anche se non strettamente attinente alle domande proposte – vedrà annullata la propria prova.
- ◆ Scrivere solo sui fogli distribuiti, cancellando le parti di brutta con un tratto di penna. Non separare questi fogli. Non utilizzare la penna rossa. Scrivere nome e cognome su tutti i fogli utilizzati.
- ◆ Tempo a disposizione: 1 ora e 45 minuti.

- 1) Si descriva come può essere partizionato lo spazio dei documenti rispetto ad una query (si illustri il significato delle diverse parti) e quali caratteristiche dovrebbe avere un SRI ideale rispetto a tale partizionamento.
- 2) Descrivere Google PageRank: cosa è, a cosa serve, su quali idee si basa e quali svantaggi ha?
- 3) Dire cosa è la ricerca semantica nei SRI web, quali sono i suoi obiettivi e descrivere il principale problema che essa deve superare.
- 4) Descrivere il processo di lavorazione “singolo input, output multipli” per la produzione editoriale basata su XML.
- 5) Dire cosa si intende per “content-based image retrieval” nell'ambito dei SRI multimediali.
- 6) Si elenchino i diversi tipi di interrogazione messi a disposizione dai sistemi per il recupero delle informazioni. Si descrivano in particolare le interrogazioni vettoriali.
- 7) Brevemente, nella pagina dei risultati di un motore di ricerca cosa è lo “snippet”?