

**Sistemi per il recupero delle informazioni**  
**Laurea Magistrale in Editoria e Giornalismo**  
**Prova scritta del 5 luglio 2013**

Cognome e nome: \_\_\_\_\_ Matricola: \_\_\_\_\_

Domanda 1	Domanda 2	Domanda 3	Domanda 4	Domanda 5	Domanda 6	Totale
/6	/5	/5	/6	/6	/5	/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.
- ◆ Tempo a disposizione: 1 ora e 45 minuti.

- 1) Disegnare e descrivere struttura, parti e funzionamento dell'architettura centralizzata crawler-indexer dei motori di ricerca sul web. È possibile per un sito web “comunicare” con il crawler? Se sì, come?
- 2) Dare lo schema e descrivere il processo d'indicizzazione nei SRI testuali e descrivere brevemente tutti i passi di cui è composto.
- 3) Si elenchino i diversi tipi di interrogazione messi a disposizione dai sistemi per il recupero delle informazioni. Si descrivano in particolare le interrogazioni booleane, illustrandone anche gli svantaggi.
- 4) 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.
- 5) Dire cosa si intende per “content-based image retrieval” nell'ambito dei SRI multimediali.
- 6) Nell'ambito dei sistemi informativi, descrivere quale è la differenza tra “dato” ed “informazione”.