

**Sistemi per il recupero delle informazioni**  
**Laurea Magistrale in Editoria e Giornalismo**  
**Prova scritta del 15 febbraio 2013**

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

Domanda 1	Domanda 2	Domanda 3	Domanda 4	Domanda 5	Domanda 6	Totale

**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) Dare lo schema e descrivere il processo d'indicizzazione nei SRI testuali e descrivere brevemente tutti i passi di cui è composto.
- 2) Descrivere cosa è, nell'ambito dei SRI testuali, un indice inverso, portando anche un esempio.
- 3) Definire cosa è il web spam (o spamdexing) e descriverne le principali tecniche.
- 4) Descrivere il processo di lavorazione “singolo input, output multipli” per la produzione editoriale basata su XML.
- 5) Dire cosa è la ricerca semantica nei SRI web, quali sono i suoi obiettivi e descrivere il principale problema che essa deve superare.
- 6) Si descriva cosa è un thesaurus e quali caratteristiche possiede. Si descriva inoltre in particolare la relazione preferenziale (o sinonimica) e le sue diverse tipologie.