

**Sistemi per il recupero delle informazioni**  
**Laurea Magistrale in Editoria e Giornalismo**  
**Prova scritta del 18 giugno 2012**

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 utilizzati.
- ◆ Tempo a disposizione: 1 ora e 45 minuti.

- 1) Si descriva cosa è un documento (riportando anche alcuni esempi) nel contesto dell'IR, come può essere classificato, e cosa e quali sono i diversi surrogati di un documento.
- 2) Si definisca cosa è l'XML, si descrivano le componenti costitutive di un documento XML riportando un piccolo esempio di documento XML e si descriva brevemente come esso viene utilizzato per la rappresentazione di documenti orientati alla narrazione.
- 3) Dire come avviene il processo di matching nei sistemi di information retrieval vettoriali.
- 4) Si parli della rappresentazione matriciale e delle diverse tipologie di matrici che possono definite ai fini dell'IR.
- 5) Si descrivano le misure di Fallout e Generalità per valutare le prestazioni dei sistemi per il recupero delle informazioni e il rapporto che le lega con Richiamo e Precisione.
- 6) Si descriva cosa è un thesaurus e quali caratteristiche possiede. Si descriva inoltre in particolare la relazione gerarchica e le sue diverse tipologie.