

Sistemi per il recupero delle informazioni
Laurea Magistrale in Editoria e Giornalismo
Prova scritta del 20 settembre 2013

Cognome e nome: _____ Matricola: _____

Domanda 1	Domanda 2	Domanda 3	Domanda 4	Domanda 5	Domanda 6	Totale
/6	/6	/6	/5	/5	/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) Nell'ambito del processo di indicizzazione di documenti testuali, cosa si intende con “stemming” e “lemmatizzazione”? Spiegare inoltre quale obiettivo ha l'uso di questi due passi.
- 2) Descrivere cosa si intende per ricerca personalizzata, quali scopi ha e su quali principi si basa. Inoltre, in questo ambito, cosa è un profilo utente?
- 3) Si descrivano le leggi di Zipf e Heaps.
- 4) Si descriva cosa si intende per “web spam” (o spamdexing)? Descrivere le principali tecniche di web spam.
- 5) Si descrivano le misure orientate all'utente per valutare le prestazioni dei sistemi per il recupero delle informazioni.
- 6) Cosa si intende per “efficacia” ed “efficienza” dei sistemi per il recupero delle informazioni?