

Homework Assignment 1

Due date: 28th of October 2013

Please note that even though the homework assignments are optional, you're still highly encouraged to answer them, as they will help you prepare for your final exam. You can work in a group of two or alone. Solutions can be dropped off at the institute's homework mailbox located on the 2nd floor, next to room 238. In that case, please make sure both your name and matriculation number is noted down. If your answers span more than one sheet, kindly staple them together. Another alternative is to send your homework via email to: elmaarry@ifis.cs.tu-bs.de

LECTURE 1: INTRODUCTION

EXERCISE 1.1

What is information retrieval concerned with? And what is the main problem it deals with?

EXERCISE 1.2

In which aspects is Web search similar and different to information retrieval?

EXERCISE 1.3

What is an information need? And how does it differ from a database SQL statement?

EXERCISE 1.4

How is the relevance notion defined in information retrieval? And what makes a document relevant or not relevant to a particular query?

EXERCISE 1.5

How does the bag model represent documents? And what's the basic idea it's built upon?

EXERCISE 1.6

What is an inverted index and what's the motivation for using it?

EXERCISE 1.7

Which Normal form is recommended in information retrieval CNF (conjunctive normal form) or DNF (Disjunctive normal form)? Why?

EXERCISE 1.8

One of the drawbacks of the Boolean retrieval model lies in the size of the returned result set. Why is the size typically difficult to control?