

Übungsblatt 9, Abgabe 13.07.2011 8 Uhr, Besprechung 13.07.12 Uhr

Aufgabe 35 XML Pfade

Ein Sammelband ist folgendermaßen aufgebaut: Die Wurzel enthält drei Elemente **front-matter**, **body** und **rear-matter**. **front-matter** enthält ein Vorwort **preface**, **body** enthält die Beiträge der Teilnehmenden als **paper** Element, wobei die Beiträge eines Tages jeweils in einem **day[X]** Element zusammengefasst sind (X ist die Nummer des Tages). **rear-matter** enthält ein Index **index**. Die Autoren aller Beiträge stehen in Vorwort und in den Beiträgen als **author** Element.

- wie sieht die Struktur vom Dokument aus?
- Mit welchem Pfad können folgende Ausdrücke erfasst werden?

- a) der Autor des Vorworts
- b) die Autoren der Beiträge
- c) der Autor des zweiten Tages
- d) alle Autor-Elemente

Aufgabe 36 structural term

Wie viele strukturelle Terme gibt es im Buch in Abbildung 10.1, geben Sie diese an. (Manning Übungsaufgabe 10.3)

Aufgabe 37 Vektorraummodell und XML

- 1) Beschreiben Sie kurz in eigenen Worten die wesentlichen Unterschiede zwischen ein Vektorraummodell für boolesches Retrieval und für XML Retrieval?
- 2) Geben Sie ein Beispiel für ein Query-Dokument Paar, dessen SIMNOMERGE (q, d) größer als 1,0 ist. (Manning Übungsaufgabe 10.11)