

Übung: Indizierung von XML Dateien

June 15, 2009

- Überlegen Sie sich ein XML-Format zur Indizierung für ihre Aufgabe: Volltext, Personen, Studiengänge
- Legen Sie ein paar Beispiel XML-Dateien an
- Schreiben Sie eine Indizierungskomponente für die Indizierung der XML Dateien
- In den XML-Dateien soll es ein Feld für den Boost des Dokumentes geben; Berücksichtigen Sie diesen bei der Suche
- Die Suche in den Dokumenten soll über die Kommandozeile möglich sein

- Logging: mit Numerierung der indizierten Dokumente, Indizierungszeitpunkt der einzelnen Dokumente
- Konfiguration (z.B. über Java Property-File): für Indexverzeichnis
- Beachten Sie, dass später ca. 40000 Dokumente indiziert werden sollen
- Beachten Sie auch das Exception-Handling