

Modulprüfung IR 3 im Studiengang Angewandte Informationswissenschaft Wiederholungsübung

Winfried Gödert / Klaus Lepsky

Institut für Informationswissenschaft
Fachhochschule Köln
Claudiusstraße 1, 50678 Köln
klaus@lepsy.de

Zusammenfassung

Die Modulprüfung wird veranstaltungsbegleitend als Hausarbeit erbracht. Sie besteht aus einem praktischen Aufgabenteil ($\frac{2}{3}$ Anteil an der MP) und einem theoretischen Aufgabenteil ($\frac{1}{3}$ Anteil an der MP). Aufgabenstellung und Anforderungen entsprechen den in den Veranstaltungen behandelten Inhalten.

Materialien

Für die Bearbeitung der Aufgabenstellung werden Ihnen zip-Dateien auf einem Server zur Verfügung gestellt (Dateinamen: „0.zip“, „1.zip“, „2.zip“ usw.), die *nach dem Entpacken* in einem Ordner folgende Dateien enthalten:

- eine Textdatei mit bibliografischen Fremddaten aus dem Bereich Sport im *RIS*-Format (Beispiel: „sport-01.txt“) für den Import in *Midos*; die Fremddaten stammen aus der Datenbank *Spolit* (Link: <http://www.bisp-datenbanken.de>);
- einen Fragebogen mit Fragestellungen zum theoretischen Hintergrund der Aufgabenstellung („theorie.pdf“).

Zu den zip-Dateien auf dem Server gelangen Sie mit folgendem Link:

<http://ixtrieve.fh-koeln.de/lehre/mp-ir3-w>

Die zehn zip-Dateien, die Sie dort vorfinden, sind von 0-9 nummeriert (s.o.). Laden Sie die zip-Datei herunter, deren Nummer der *letzten Stelle* Ihrer Matrikelnummer entspricht, d. h. bei der Matrikelnummer „110 12345“ laden Sie die zip-Datei mit der Nummer „5“ herunter. Die Übereinstimmung von Matrikelnummer und Dateinummer ist *Bedingung für die gültige Abgabe*. Ausarbeitungen mit fehlender Übereinstimmung werden nicht bewertet.

Teilaufgaben

Praktischer Teil (2/3 Anteil)

1. Import der bibliografischen Daten aus der entsprechenden Textdatei in eine neue *Midos*-Datenbank unter Erstellung bzw. Anpassung eines geeigneten Kategorienschemas und Verwendung eines Texteditors zur Vorbereitung der Textdatei für den Import;
2. Erstellung von mindestens zwei Thesauri mit *Midos*-Thesaurus zur intellektuellen Erschließung im Rahmen eines selbst gewählten *aspektorientierten Erschließungskonzepts*;
3. Intellektuelle Erschließung von *20 Datensätzen* gemäß des selbst gewählten aspektorientierten Erschließungskonzepts mit den Thesauri;
4. Erstellung von *mindestens* zwei montierten Ausgabeformaten:
 - ein Ausgabeformat, das *nur* die formalen Merkmale der Datensätze anzeigt;
 - ein Ausgabeformat, das die formalen, *alle* inhaltlichen Merkmale *und* die Lingo-Daten zu den Datensätzen anzeigt;
5. Erstellung einer Retrievalanwendung mit mindestens folgenden Eigenschaften:
 - Navigation über Frame,
 - Suchmaske mit Suchindizes und mouse-over Hints,
 - Thesaurus-Einbindung aller Thesauri,
 - Verlinkung geeigneter Kategorien für die direkte Navigation in der Vollanzeige,
 - Startbildschirm mit Hinweisen zur Benutzung der Retrievalanwendung,
 - Einbindung *aller* Ausgabeformate;
6. Automatische Indexierung der inhaltlich relevanten Kategorieninhalte aller Datensätze mit *Lingo* mit mindestens folgendem Funktionsumfang:
 - Lemmatisierung,
 - Dekomposition,
 - algorithmische Mehrworterkennung mit den Erkennungsmustern „AS“ und „AK“;Achtung: Speichern Sie die Protokolldatei des Indexierungslaufs mit der Endung „*.log“ auf der CD ab!
7. Import der Ergebnisse der automatischen Indexierung in die *Midos*-Datenbank in zwei getrennte Kategorien:
 - eine Kategorie für alle Grundformen, die identifizierten Komposita und deren Teilwörter;
 - eine Kategorie für die Ergebnisse der algorithmischen Mehrworterkennung (sequences);
8. Erweiterung der Retrievalanwendung um Such- und Anzeigeoptionen für die Ergebnisse der automatischen Indexierung.

Theoretischer Teil (1/3 Anteil)

- Beantwortung der theoretischen Fragen („theorie.pdf“); Die Bearbeitung der Wiederholungsfragen bietet eine gute Vorbereitung auf den theoretischen Prüfungsteil.

Bewertungskriterien

- Vollständigkeit und prinzipielle Sinnhaftigkeit der Aufgabenerfüllung;
- Vollständigkeit des Imports in die Datenbank;
- Angemessenheit des Kategorienschemas für die Dokumentbeschreibungen;
- Stimmigkeit des aspektorientierten Erschließungskonzepts;
- Stimmigkeit der sprachlichen Gestalt der Deskriptoren in den Thesauri;
- Vorhandensein aller Relationsarten (Synonyme; Oberbegriffe, Unterbegriffe; Verwandte Begriffe) sowie einiger Definitionen in den Thesauri;
- Einbindung der Deskriptoren in das hierarchische Gefüge;
- Angemessenheit der inhaltlichen Beschreibung bzw. intellektuellen Erschließung der Dokumente;
- Vollständigkeit der Ergebnisse der automatischen Indexierung mit *Lingo*;
- Inhalt der Datei „*.log“;
- Tauglichkeit der erstellten Retrievalumgebung für formale und inhaltliche Recherchen;
- Aussagekraft der Beschreibung der Datenbank und ihrer Suchfunktionen auf dem Startbildschirm;
- Korrektheit der Antworten im theoretischen Teil.

Regularien

- Einzelarbeit;
- Abgabe der CD (Archivexemplar) gemeinsam mit dem ausgefüllten Fragebogen zum theoretischen Teil spätestens zum angegebenen Termin im Postfach *Gödert* des Instituts (Zusendungen per Email werden nicht akzeptiert!);
- Bewertungsgegenstand sind allein die Retrievalanwendung auf der CD und der ausgefüllte Fragebogen zum theoretischen Teil;
- Beschriftung der CD: Name, Matrikel-Nr., Studiengang/Semestergruppe, Email-Adresse; (diese Angaben bitte ein zweites Mal in einer Textdatei auf der CD hinterlegen).

Beginn der Bearbeitung: erfolgt selbstbestimmt

Abgabe der Ergebnisse: