

Kurzfassung der Diplomarbeit

Abteilung	Informatik
Name der Studenten	Reto Senn Ulrich Meierhofer
Diplomjahr	2004
Titel der Diplomarbeit	Glossar- und Thesaurus-Editor mit Eclipse
Examinatorin / Examinator	Claude Eisenhut, Eisenhut Informatik Jegenstorf b. Bern

Kurzfassung der Diplomarbeit

Der Glossar- und Thesaurus-Editor (T-Saurus) erlaubt es, schnell und einfach Begriffe zu definieren, zu klassieren und untereinander zu verknüpfen. Der Editor ist so konzipiert, dass er einem Benutzer erlaubt, eine "Begriffsdatenbank" (u.a. mit Synonymen, Unter- und Oberbegriffe) zu erstellen und diese z.B. auch für die Unterstützung von spezialisierten Suchdiensten zu verwenden.

Auf dem Markt sind schon einige Programme zur Erstellung von Thesauri erhältlich. Keines deckt jedoch unsere Bedürfnisse einer kompletten Lösung ab: Entweder ist die Software veraltet, umständlich zu bedienen, nur auf ein gewisses Thema zugeschnitten oder nicht Open Source; ausserdem gibt es nur wenige Lösungen, welche es erlauben, Glossar und Thesaurus gleichzeitig zu erstellen.

Ziel unserer Arbeit war ein einfach zu benutzender Editor, welcher möglichst universell einsetzbar ist. Das heisst u.a.: Benutzerfreundliches GUI, automatisches Verknüpfen von Begriffen und Unabhängigkeit von einem Themengebiet. Zudem sollte der Editor auf der Eclipse-Plattform aufbauen, also als Komponenten-Gruppe (Plug-in) entwickelt werden. Neben den verschiedenen Vorteilen dieser Komponenten-Architektur haben wir uns davon vor allem eine kürzere Entwicklungszeit, bzw. mehr Funktionalität versprochen, da wir so bereits vorhandene Komponenten nutzen können. Damit wird auch die Erweiterbarkeit des Editors verbessert.

Die Architektur der Daten ist logisch und einfach. Die internen Daten sind mit einem XML Schema definiert und dokumentiert. Sie können aus standardisierten Formaten (z.B. CSV) importiert und in standardisierte (z.B. DocBook) oder proprietäre Formate exportiert werden. Entsprechende Filter werden dynamisch geladen und angesprochen - bei Bedarf können damit beliebig weitere Formate unterstützt werden.