



TECHNISCHE UNIVERSITÄT WIEN

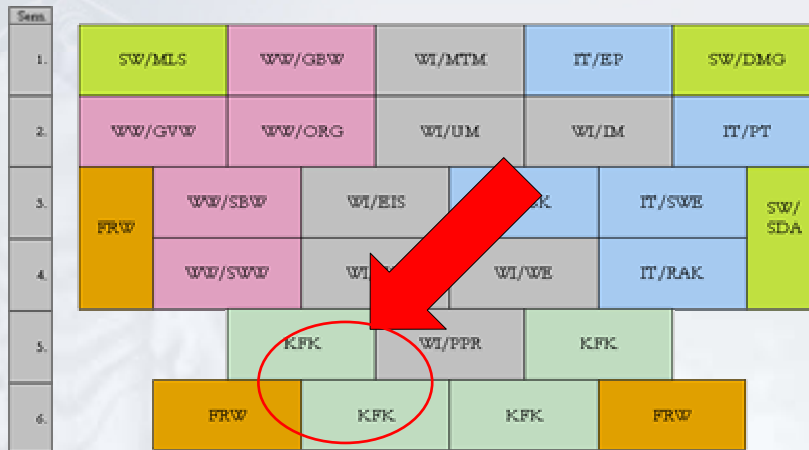
Semantic Web Praktikum ..und andere Praktika... WS 2004/05

*Robert Baumgartner, Jürgen Dorn,
Georg Gottlob, Marcus Herzog*



KfK Semantic Web

- Kernfachkombination
 - Wirtschaftsinformatik
 - Vertiefendes Wahlfach
 - [DBAI \(http://www.dbai.tuwien.ac.at\)](http://www.dbai.tuwien.ac.at) und [DKE \(http://www.dke.univie.ac.at\)](http://www.dke.univie.ac.at)
- Module der KfK
 - Einführung in das Semantic Web
 - Advanced Topics
 - Knowledge Management
 - **[Semantic Web Praktikum](#)**

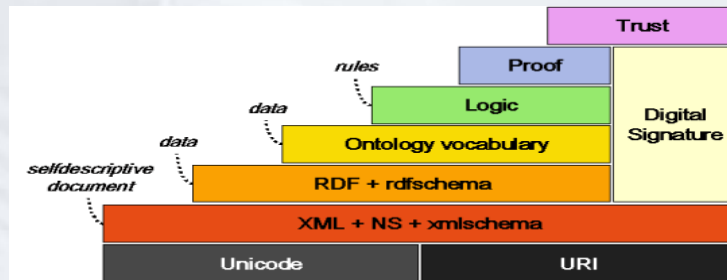


Organisation des Studienablaufs des Bakkalaureatsstudiums Wirtschaftsinformatik

Semantic Web Praktikum

- 4 Wochenstunden
 - Ingesamter Aufwand ca. 60 Stunden
- Plus: Bakkalaureatsarbeit
 - (wenn für KfK Bakk. gewählt)
 - Seitenanzahl ca. 30 Seiten
- Themenbesprechung individuell
 - Beginn jederzeit möglich, nicht nur am Semesteranfang
- Auch Arbeiten in Kleingruppen möglich
 - Vorzugsweise aber mit eigener Bakkalaureatsarbeit

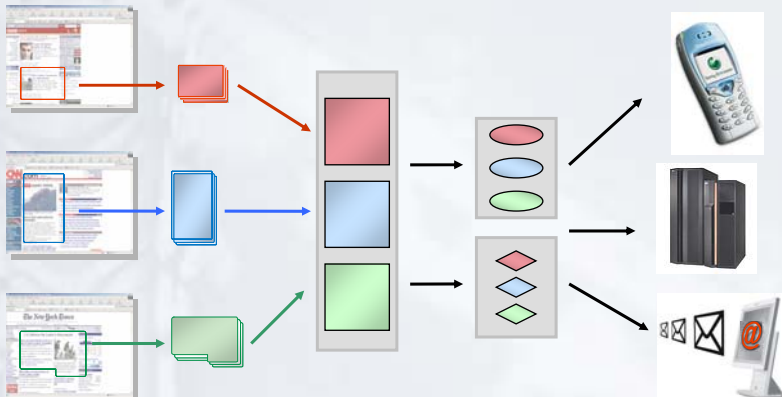
Motivation – Semantic Web



Quell Publikation Extraktion Integration Transformation Ziel Publikation

Lixto Visual Wrapper

Lixto Transformation Server



HTML

XML

WML, XHTML, Text

Themenbereiche

- Web-Datenextraktion und Semantic Web
 - Robert Baumgartner, Marcus Herzog
- Ontologien, KI Komponenten, Semantic Web, Web Services
 - Jürgen Dorn

Themenauswahl – Webdatenextraktion

- Applikationserstellung mit Lixto
 - Lixto Suite Beispielapplikationen, z.B. im Bereich Business und Competitive Intelligence, Portalintegration, oder auch tägliche Speisepläne der Restaurants in der Umgebung
- Vergleich Lixto Suite mit anderen Wrapper-Generatoren
 - Vergleich/Erweiterung mit/von RISE Wrapper Repository unter <http://www.isi.edu/info-agents/RISE/>
 - Studie über einige andere Tools insb. WebQL, NQL
- Applikationen mit Mozilla XUL
 - <http://books.mozdev.org/chapters>
 - Zugriff und Manipulation DOM Baum
 - Kommunikation mit (Ontologie-)Datenbanken

Themenauswahl – XML und Semantic Web

- RDF Repositories im Vergleich
 - Semantische Abfragesprachen
 - Ontologie/Taxonomie-Modellierung
 - Sesame, Jena, KAON, TripleStore
- RDF/OWL Repositories for Lixto
 - Füllen von Ontologien mit Lixto Instanzen
- RDFS Taxonomien aus XML Schemata
 - Spezifikation/Implementation semiautomatische interaktive Abbildung

Themenauswahl – Visual Wrapper Implementierung

- Wrapper Adaptation
 - Site Statistiken
- Mitarbeit bei Extraktion aus anderen Formaten
 - PDF
 - Textuelle Daten
- Neue Beispiele und Verbesserung der Lixto Visual Wrapper Regression Testsuite
 - Testwrappererstellung
 - Arbeiten mit ANT

Themenauswahl – XML Trafos und Workflows

- Graphische Datentransformations- und abfragesprachen/tools im Vergleich
 - BizTalk Mapper, Tibco, ...
 - XCept, XML-GL, ...
- BPEL vs. DAML-S/OWL-S
 - Und Vergleich Workflowstandards und Engines
- Updating XML Structures / Diffing XML Structures
- Study useful Eclipse Plugins
 - Graph Component etc.
- Lixto und andere Werkzeuge
 - Portalgenerierung: Intrexx Portalmanager
 - Klassifizierung: L4
 - Text Annotation mit GATE (<http://gate.ac.uk/ie>) als Hilfe zur Webinformationsextraktion in Lixto
- Gelegentlich gibt es auch Praktikumsthemen in Zusammenarbeit mit Firmen

Ressourcenmanagement

- Entwicklung eines Mitgliederverwaltungssystems für Mitglieder einer Organisation
 - LDAP Server und MySQL DB
 - XForms basierte Benutzerschnittstelle, Abfrage mit Web-Service
- Entwicklung eines Bücher und Zeitschriftenverwaltungssystems
 - MySQL DB
 - XForms basierte Benutzerschnittstelle, Abfrage mit Web-Service
- Entwicklung eines Verwaltungssystems für Räume
 - MySQL DB
 - XForms basierte Benutzerschnittstelle, Abfrage mit Web-Service

Ontologien

- Transformation der OTA Terminologie (Nachrichten) für Tourismus in eine OWL Ontologie
 - Protege Open Source Ontologieeditor
 - Referenzmodell liegt möglicherweise nur in Papierform vor
- Transformation der RMSIG Ontologie für Tourismus nach OWL
 - Protege Open Source Ontologieeditor
 - Referenzmodell liegt möglicherweise nur in Papierform vor
- Definition einer Office-Ontologie in OWL
 - Ressourcen (Räume, Bücher, ...), Personen,

Integration von KI-Komponenten

- Integration eines JESS Moduls als Web Service
 - Zugriff von JESS auf Datenbank und/oder Ontologie
 - kleines Beispiel mit Vorwärtsverkettung
- Integration eines Prolog Moduls als Web Service
 - Zugriff von Prolog auf Datenbank und/oder Ontologie
 - kleines Beispiel mit Rückwärtsverkettung
- Integration eines Racer Moduls als Web Service
 - Zugriff von Racer auf Datenbank und/oder Ontologie
 - kleines Beispiel

Serverbasierte Lösung zur Verwaltung von persönlichen Daten

- vCard Server mit Web Service Schnittstellen
 - Eingabe über XForms Schnittstelle
 - Zugriff auch über Handy
- vCalendar Server mit Web Service Schnittstellen
 - Eingabe über XForms Schnittstelle
 - Zugriff auch über Handy

Entwicklung eines Umfragewerkzeugs

- Evaluierung existierender Programme
- Einsatz von XForms
- verschiedene Kanäle (Internet, Papier, Handy)

Erzeugung von Wissensbasen mit Lixto

- Extraktion von Daten aus Wiener Kultur Web Seiten
 - Definition einer Ontologie
 - Ablage der Daten in MySQL DB
 - Realisierung eines „Agenten“, der über Web Services aus der Wissensbasis Fragen beantwortet
- andere Domänen auch möglich

Kontaktdaten der Vortragenden

- Robert Baumgartner
 - <http://www.dbai.tuwien.ac.at/staff/baumgart>
 - baumgart@dbai.tuwien.ac.at
- Jürgen Dorn
 - <http://www.dbai.tuwien.ac.at/staff/dorn>
 - dorn@dbai.tuwien.ac.at
- Georg Gottlob
 - <http://www.dbai.tuwien.ac.at/staff/gottlob>
 - gottlob@dbai.tuwien.ac.at
- Marcus Herzog
 - <http://www.dbai.tuwien.ac.at/staff/herzog>
 - herzog@dbai.tuwien.ac.at