ALPEN-ADRIA
UNIVERSITÄT
KLAGENFURT

EINLADUNG

Informatik-Sonderkolloquium

VORTRAGENDER: **AO. PROF. DR. WERNER RETSCHITZEGGER,**
JOHANNES-KEPLER-UNIVERSITÄT LINZ,
ABTEILUNG FÜR INFORMATIONSSYSTEME

TERMIN: DONNERSTAG, 17. NOVEMBER 2005

ORT / ZEIT: HÖRSAAL HS 3, 16.15 UHR

TOWARDS INTEGRATING HETEROGENEOUS DATA SOURCES

The nature and sources of data are constantly changing, not least due to new technologies such as the Web and upcoming engineering principles, such as model-driven development. Looking at the area of *Web Engineering*, the *Extensible Markup Language (XML)* is generally accepted as data description language for both web-based information systems and electronic data interchange between different organizations. Database Systems (DBS) are often used as a backbone for the consistent storage, retrieval, and manipulation of XML data, thus requiring approaches for seamless integration. Apart from the area of Web Engineering, with the rise of *Model Engineering*, more and more development tasks shall be performed on models, replacing code as the major artifact in software development. Seamless exchange of models among different modeling tools increasingly becomes a crucial prerequisite for effective software development processes. Consequently, no matter whether in the area of Web Engineering or in the area of Model Engineering, appropriate facilities for the integration of heterogeneous data sources are more than ever of paramount importance. This talk will focus on three integration challenges being routed in these application areas.

First, the *efficient generation and maintenance of XML-based web pages out of a relational DBS* is investigated. A novel approach for *self maintaining web pages (SMWP)* is presented that is, different to previous approaches, characterized by a simple and thus easy to maintain mapping between web pages and databases. This is achieved by utilizing fragmentation techniques from distributed databases, by allocating parameterized fragments to web page classes and by employing an incremental push based data delivery approach.

The second part of the talk focuses on the role of XML for electronic data interchange and proposes a *flexible mapping between XML and relational databases.* The key idea of the approach called *X-Ray (Integrating XML and Relational Database Systems)* is that mappings may be defined between XML DTDs and relational schemata while preserving their autonomy. This is made possible by introducing a meta schema and meta knowledge for resolving data model heterogeneity and schema heterogeneity. Since the mapping knowledge is not hard-coded but rather reified within the meta schema, adaptability to schema changes is enhanced. The meta schema provides the basis for X-Ray to automatically compose XML documents out of the relational database when requested and decompose them when they have to be stored.

The final part of the talk sketches *ModelCVS, a vision for a semantic infrastructure aiming at model-based tool integration*. ModelCVS enables transparent transformation and versioning of heterogeneous models, going well beyond existing low-level model transformation approaches. For this, ModelCVS utilizes semantic technologies and integration patterns, allowing integration at the metamodel level. To foster reuse, a repository captures essential knowledge about modeling languages and tools in terms of ontologies.

Wrapping up, we will reflect on the practical relevance of the integration mechanisms in use and point to open research questions.