

Bemerkungen zu einem Datenmodell der Zukunft

HTWK 20.05.2015
15.30 - 17.00
Gustav Freitag Straße 42a
Gutenbergbau 112

Dr. habil. Klaus Benecke
(benecke-systeme)

Datenstrukturen mit Wiederholgruppen gibt es in der Informatik bereits seit mehr als 50 Jahren. Im Relationalen Datenmodell wurden sie eliminiert und mit XML (*,+) und JSON, ... sind sie wieder zurück. Im Vortrag wird deutlich, dass ein Datenmodell nicht nur für Informatiker von Bedeutung ist, sondern für alle Computernutzer, die Informationssysteme benötigen. Wir brauchen ein Datenmodell, das über so wenige wie möglich, leicht nutzbare Konstrukte und Werkzeuge verfügt. Auch einfache elementare Datentypen wie WORT und STRICH, werden in Zukunft bedeutungsvoll sein. Im Vortrag wird eine abstrakte Präzisierung von XML-Dokumenten angegeben, die zwischen Tupel und Kollektion unterscheidet. Da diese Strukturen Tabellen und Dokumente umfassen, nennen wir sie TABMENTe. Auf der Basis von Tabmenten lassen sich leistungsfähige, natürliche Operationen definieren und implementieren. Neben der Umstrukturierung (**gib**) (restrukturieren, aggregieren, Duplikate eliminieren, vereinigen, ...) werden die Selektion (**mit, ohn**), Extension (**ext**) (Berechnungen, und Join) und ein noch nicht implementiertes einfach zu handhabendes, natürliches Konstrukt **ext+** und ein Konstrukt **gib+** kurz vorgestellt.

Ferner wird deutlich, dass einfache algebraische Optimierungsregeln für die Anfrageverarbeitung nicht mehr universell gelten (Kommutativität von Bedingungen).

Neben Anfragen an ganze Datenbanken werden mehrere Anfragen an die deutsche Wikipedia vorgestellt. Die Wikipedia ist ein sehr schönes Anwendungsgebiet für Tabmente, da ihre Objekte Dokument- und Tabellenanteile enthalten. Wikidokumente werden in der Regel nicht als XML-Dokumente aufgefasst; sie sind jedoch im Tabmentbegriff eingeschlossen. Im Vortrag soll deutlich werden, dass die Wikipedia ein großer Wissensschatz ist, der mit einer leistungsfähigen Anfragesprache weit besser zu heben ist. Ferner zielt der Vortrag darauf, Studenten und Absolventen zu gewinnen, die sich an Teilprojekten von OttoPS in Form von Praktikums-, Bachelor- oder Masterarbeiten oder im Rahmen von EXIST- oder anderen Maßnahmen beteiligen. Das bisherige Pilotprojekt **OttoPS** ist in der funktionalen Programmiersprache **OCAML** implementiert. Dokumentationen und eine Version zum Testen findet man unter **ottoPS.eu**.