

Graduiertenkolloquium Angewandte Informatik

Ranking for Heterogenous Web Data Search

Dipl.-Wi.-Ing. Daniel M. Herzig
AIFB

Abstract:

Ranking - die algorithmische Entscheidung, wie relevant ein Artefakt für einen gegebenen Informationsbedarf ist - ist ein integraler Bestandteil jeder Suchmaschine, der deren Effektivität maßgeblich beeinflusst und dafür sorgt, dass trotz zunehmender Datenmengen die relevanten Suchergebnisse gefunden werden. Neben den textuellen Daten im Web werden durch die Anstrengungen der Semantic Web Community und zunehmend durch die großen Suchmaschinen immer mehr strukturierte Daten im Web veröffentlicht, die neue Möglichkeiten für Suche im Web eröffnen. Diese strukturierten Daten, die größten Teils den Linked Data Prinzipien folgen, beschreiben und verlinken Entitäten und erlauben so dem Benutzer faktenbasierte Suchergebnisse anstelle von Links zu präsentieren. Da viele Datenquellen ähnliche oder die gleichen Entitäten, wie z.B. Produkte, Filme, Bücher oder Orte, beschreiben, diese jedoch unterschiedlich modellieren, entsteht eine heterogene Datenlandschaft, die es Anwendungen erschwert diese Daten nutzbar zu machen. Selbst wenn die Heterogenität der Daten überbrückt wurde, schließt sich das Problem der Redundanz an, da die Inhalte der Datenquellen sich ergänzen, aber teilweise auch überschneiden und so mehrfach Beschreibungen für dieselbe Entität vorkommen. Beide Herausforderungen, Heterogenität und Redundanz, stellen Probleme für die Nutzung von strukturierten Webdaten für Suchmaschinen und insbesondere deren Rankingmethoden dar. In diesem Vortrag werden die zwei Probleme - Heterogenität und Redundanz - von Webdaten erörtert und eine Rankingstrategie für strukturierte Webdaten vorgestellt, die diese Probleme berücksichtigt und gegenüber bestehenden Ansätzen eine effektivere Suche ermöglicht.

Termin: Mittwoch, 06.03.2013 15:45 Uhr

Ort: Englerstraße 11, 76131 Karlsruhe
Kollegiengebäude am Ehrenhof (Geb. 11.40), 2. OG, Raum 231
(Hinweise für Besucher: www.aifb.kit.edu/Allgemeines/Besucher)

Veranstalter: Institut AIFB, Wissensmanagement

Zu diesem Vortrag lädt das Institut für Angewandte Informatik und Formale Beschreibungsverfahren alle Interessierten herzlich ein.

Andreas Oberweis, Hartmut Schmeck, Detlef Seese, Wolffried Stucky, Rudi Studer (Org.), Stefan Tai