

Beispiele für semantische Aufbereitung

Semantic-Web-Lösungen dienen dazu, Daten mit mehr Informationen anzureichern und damit leichter nutzbar zu machen. Das geschieht mit Hilfe von Metadaten (Daten über Daten), die einem Objekt (der Information) zugewiesen werden. Das Semantic Web ist noch in der Entwicklung, die Beispiele sind alle nicht marktreif. Die Grundlagen (z.B. Ontologien, Abfragesprachen) sind inzwischen jedoch größtenteils verfügbar. Die Daten des bisherigen Webs müssten jedoch für das Semantic Web neu aufbereitet werden. Ob das gelingt und sich durchsetzt, ist umstritten.

Im folgenden einige Beispiele, in denen semantischer Analyse eingesetzt wird:

Dandelon

Folgender Algorithmus läuft ab: Es werden von deutschen orten die Grundwortformen ermittelt (Häusern-> Haus), im Englischen greift Stemming, die Syntax der Query wird geprüft und Mehrwortbegriffe, die als solche nicht vorkommen (wie "Haus Kunst"), werden mit einem logischen UND verknüpft. Die einzelnen Begriffe werden in den 1.26 Millionen Thesauruseinträgen gesucht (mehrere Datenbanken) und per default werden Synonyme, Übersetzungen und Grundwortformen an die Query angehängt (More Precise). Findet diese Suche noch kein Ergebnis im Datenbestand, werden automatisch Unterbegriffe ergänzt (Medium). Führt dies noch immer nicht zum Ziel wird die Wortstammsuche (*stamm*) oder Fuzzy-Search ausgeführt (More Recall). Wer damit noch nicht fündig wurde, kann selbst aktiv werden und über die grafische Navigation (Topic Maps Visualization) weitere Begriffsassoziationen evtl. finden und als neue Query nutzen. Die Ergebnisanzeige ist per default nach Ranking sortiert (alternativ Jahr oder Autor), Ranking-Werte unter 70 deuten darauf hin, dass die gefundenen Werke nicht als sehr einschlägig erkannt wurden. Das Highlighting in Titel und Schlagworten betont optisch eine mögliche Relevanz.

<http://www.dandelon.com/intelligentSEARCH.nsf/fmQSF?OpenForm>

Conweaver Bauportal

Diese Probleme soll die semantische Suchmaschine ConWeaver beseitigen. An Stelle eines Volltextindexes setzt sie ein Wissensnetz, auch semantisches Netz oder Ontologie genannt, als reichhaltigen wissensbasierten Suchindex ein. Semantische Wissensnetze bilden Inhalte als Netz aus Begriffen, logischen Begriffsbeziehungen und vielfältigen Begriffsbezeichnungen ab. ConWeaver baut dieses Wissensnetz im Unterschied zu den meisten anderen ontologiebasierten Softwareprodukten weitgehend automatisch auf und macht dadurch semantische Technologien auch fuer die Unternehmen erschwinglich, die bisher aus Kostengründen davor zurueckgeschreckt sind. ConWeaver liefert in mehrsprachigen, heterogenen Datenumgebungen mit einer Vielfalt von Dokumenttypen - wie sie typischerweise in Intranets, Unternehmensportalen, Enterprise-Content-Management- oder Lotus-Notes-Systemen vorliegen - weitgehend vollstaendige und ausreichend praezise Suchergebnisse, weil die Suchmaschine Bezeichnungsalternativen, Formulierungsvarianten, Uebersetzungen oder auch falsch geschriebene Begriffe zuverlaessig einer Bedeutung zuordnet. Zudem integriert ConWeaver hochstrukturierte Fakteninformationen aus relationalen Datenbanken und unstrukturierte Informationen aus Textdokumenten und fasst sie in einem inhaltlich aggregierten und thematisch strukturierten Suchergebnis zusammen, statt nur eine reine Dokumentliste auszugeben. Dabei bleibt die Bedienung angeblich so einfach wie in Google, das Ergebnis ist jedoch aehnlich praezise wie eine Datenbankabfrage.

<http://www.conweaver.de>

<http://www.ipsi.fraunhofer.de/contec>

(die Software bleibt vorerst Unternehmen vorbehalten, die sie von den Fraunhofer-Wissenschaftlern in ihren Intranets installieren lassen).

APA Online Manager V 6

Werkzeuge zur Medienbeobachtung mit semantischer Technologie

<http://www.buzinkay.net/blog-de/2006/07/semantische-tools-der-apa-online-manager/>

neueste Technologien aus den Bereichen

Banking Intelligence Projekt DFKI/VW Financial Services

Semantic Web und Semantic Desktop fließen in die Entwicklungsarbeit mit ein. Diese ermöglichen den ungehinderten Austausch vorgehaltener Datenbestände über (weltweit) verteilte Arbeitsplätze und über unterschiedliche Anwenderprogramme hinweg. Durch die proaktive Bereitstellung relevanter Informationen ermöglichen sie gleichzeitig die zielführende Strukturierung wissensintensiver Arbeitsprozesse.

<http://www.ita-kl.de/ita/forschung/projekt.php?navid=32&projektid=930&engl=1>

Projekt Reisewissen (FU Berlin)

<http://reisewissen.ag-nbi.de>

LCE-Guide/BINE-Projekt

Das neue BINE-Projekt-Info "Produkte energieeffizient gestalten" stellt die Internet-basierte Wissensdatenbank LCE- Guide vor . Vorwiegend die Elektro- und Automobilbranche wird hier ganzheitliche Informationen zu technischen, wirtschaftlichen und ökologischen Kriterien von Werkstoffen und Bauteilen abrufen können. Um bereits während der Entwicklung eines Produkts die Auswirkungen auf alle Lebensphasen und Prozesse abschätzen zu können (Life Cycle Design), sind viele Informationen notwendig. Die Internet-basierte Plattform, die dieses Wissen bereitstellt und gleichzeitig andere Aspekte wie Recht und Wirtschaftlichkeit mit abdeckt, soll eine an praxisgerechte Entscheidungshilfe bieten und den Einsatz innovativer Materialien und energieeffizienter Techniken fördern. Zur Zeit unterziehen verschiedene Industrieunternehmen den LCE-Guide einem Praxistest.

<http://www.bine.info>

www.lce-guide.de

Kartoo Die Suchmaschine Kartoo zeigt eine kartografische Aufbereitung der Ergebnisliste. Die thematischen Relationen zwischen den Ergebnissen werden dabei mit Verbindungslinien angezeigt. Kartoo bezieht dabei u.a. die Suchmaschinen t-online, Tricus, Alta Vista, Abacho, Fireball, Yahoo und MSN mit ein.

<http://www.kartoo.com>

(frei nutzbar)