

Neue Geschäftsmodelle möglich _Semantische Suche

Die verlegerischen Vorzüge des vernetzten Wissens

Je größer das Informationsvolumen, desto wichtiger die Navigation: Der Gentner Verlag hat eine semantische Suchfunktion implementiert.

Die Autorin Marion Fugléwicz-Bren ist freie Autorin, Journalistin und Communications Consultant in Wien. Seit 2007 arbeitet die Autorin auch für die Semantic Web Company (www.semantic-web.at).

Wer kennt nicht die Verzweiflung des Google-Suchers, der nach einem speziellen Thema forscht? Die Fülle irrelevanter Information kostet Zeit und Nerven. Maschinen, die „verstehen“, wofür sich der Suchende interessiert, können zielgenauer Resultate liefern. Das Semantic Web soll das Internet leichter navigierbar machen, die semantische Suche dabei helfen, die Inhalte zu finden, die man tatsächlich sucht, selbst wenn man es noch gar nicht weiß. Was in manchem Ohr noch wie ein Märchen klingt, daran wird schon seit Jahren fieberhaft gearbeitet, auch in der Buch- und Zeitschriftenbranche. Das Konzept dahinter besteht vor allem darin, Zusammenhänge sichtbar zu machen, damit – an verschiedenen Stellen gespeicherte – Information als vernetztes Wissen zur Verfügung steht.

Vorteile der semantischen Suche für Verlage

Eine semantische Suche geht weit über die Volltextsuche hinaus, sie berücksichtigt nicht nur Schlagworte, sondern die Ähnlichkeit gesamt

er Texte. Sie ist somit für die Redaktion ein wertvolles Recherche-Werkzeug. „Zielgruppen und ihre Interessen werden immer feiner, dem muss auch bei der Lieferung von Content Rechnung getragen werden“, erzählt der Informatiker und Sologe Torsten Weise. Als Berater für mittelständische Fachverlage liegt sein Schwerpunkt auf der produkt-

orientierten Integration neuer Technologien. Für den Gentner Verlag begleitete Weise die Implementierung einer semantischen Suchfunktion.

Auch für kleine Verlage einsetzbar

Gentner verfügt laut Weise über einen großen Korpus von Fachinformationen, der jedoch jenseits der Zuordnung zu Rubriken im Heft und der Volltextsuche nicht nutzerfreundlich angeboten werden konnte. „Es gab hier ein erstklassiges inhaltliches Angebot, dessen Content jedoch nicht einfach aufzuspüren oder zu nutzen war.“ Diese mangelhafte Sichtbarkeit sollte behoben werden: Die zu ladenden Artikel wurden im Content-Management-System identifiziert und zusammen mit externen Word-, Excel- und PDF-Dateien „indiziert“. Eine – für spezifische Benutzer optimierte – Oberfläche wurde für die Suche entwickelt. Insgesamt hat der Prozess etwa drei Monate gedauert.

Ein solcher Schritt ist nicht notwendigerweise eine Frage der Größe eines Verlags, so Weise: „Wesentlich ist der Umfang der Inhalte und insbesondere das Potenzial, die semantische Suche beziehungsweise deren Ergebnisse für neue Produkte, Angebote oder Kanäle zu nutzen.“

Primär im Fokus der Fachverlage

Auf die Frage, ob solch eine Implementierung nur für Fachverlage sinnvoll sei, erklärt Weise: „Kern der semantischen Suche ist es, die thematische, linguistische Ähnlichkeit von Inhalten mit statistischen Methoden zu bestimmen, wobei die Inhalte durchaus als PDF, Power-Point-Präsentation oder in Tabellenform vorliegen können. Bei literarischen Texten halte ich den Nutzen deshalb für begrenzt. Anbieter von so genannten

Gentner Verlag

Aus dem Stuttgarter Stammhaus des Gentner Verlages sind über die Jahre verlegerische Aktivitäten mit knapp 40 Fachzeitschriften entstanden. Die Schwerpunkte sind die Stammgebiete Sanitär- und Heizungstechnik, Kälte- und Klimatechnik, Glas sowie Elektro- und Sicherheitstechnik. Die Redakteure der Branchenzeitschrift „SBZ – Sanitär, Heizung, Klima“ verfügen selbst über handwerkliche Ausbildungen und haben in den Branchen gearbeitet, über die sie schreiben. Das zweite Branchenstandbein des Verlags ist der Medizinbereich, dazu gibt es vier Fachzeitschriften. www.gentner.de

Special-Interest-Publikationen können hingegen sehr gut von der semantischen Suche profitieren.“ Grundsätzlich gelte: Je größer der Korpus, desto bessere Ergebnisse der semantischen Suche.

Automatische Ableitung aus dem Kontext

Der Charme der semantischen Suche bestehe darin, dass die meisten thematisch verwandten Begriffe sich automatisch aus dem Kontext ableiten ließen: Das System erkenne, dass zwei Begriffe häufig zusammen erwähnt werden – in einem Satz, einem Absatz, einem Dokument – und folgere daraus, dass sie miteinander in Bezug stehen. Ein passendes Dokument könne also selbst dann gefunden werden, wenn der Suchbegriff selbst nicht enthalten sei. Zu diesem Zweck müsse also nicht erst ein Thesaurus aufgebaut werden, anders als in vielen auf dem Markt erhältlichen Lösungen würden auch keine vorgefertigten Ontologien verwendet. Dennoch gebe es Fälle, bei denen dem System noch geholfen werden müsse. Beispiel: Die Begriffe „Klempner“, „Flaschner“ und „Blechner“ bezeichneten alle den gleichen Beruf, würden allerdings meist nur regional verwendet. Hierfür könne festgelegt werden, dass die Suche für eine dieser Bezeichnungen gleich auf die beiden anderen übertragen wird.

Automatische Ableitung aus dem Kontext

Interessant ist nach Einschätzung von Weise auch die Funktion, die mittels semantischer Suche gefundenen Artikel zusammen in ein einziges Themen-Dossier-PDF konvertieren zu können: „Gerade im wissenschaftlichen oder auch im Special-Interest-Bereich sind damit einige kostenpflichtige Angebote denkbar.“ Ein weiterer Vorteil: Durch die bessere interne Verlinkung werden Inhalte auch durch die Suchmaschinen als relevanter bewertet und erhalten somit ein besseres Ranking.

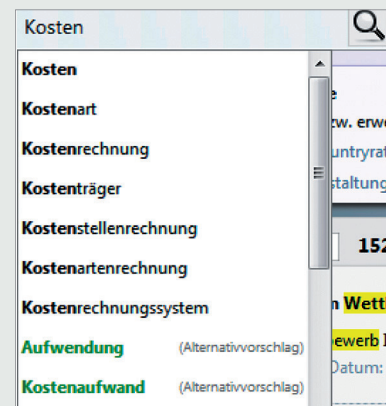
Apropos Suche: „Wir suchen niemals die Dinge, sondern das Suchen nach ihnen“, meinte einst Blaise Pascal. Auch hierbei kann die semantische Suche eine große Hilfe sein, wird man sich doch gar nicht selten erst beim Suchen bewusst, wonach man eigentlich gar nicht gesucht hatte.

Marion Fugléwicz-Bren, redaktion@buchreport.de
www.gentner.de

So funktioniert die semantische Suche

1. Automatische Vervollständigung der Suchanfragen

Bei der Eingabe eines Suchbegriffs wird eine Auswahlliste angeboten, aus der Begriffe gewählt werden können, ähnlich wie bei Google oder anderen Webdiensten. In dieser Liste werden Schlagwörter aus dem Suchindex angezeigt, zusätzlich aber auch mögliche Synonyme als Alternativvorschläge (in Grün).

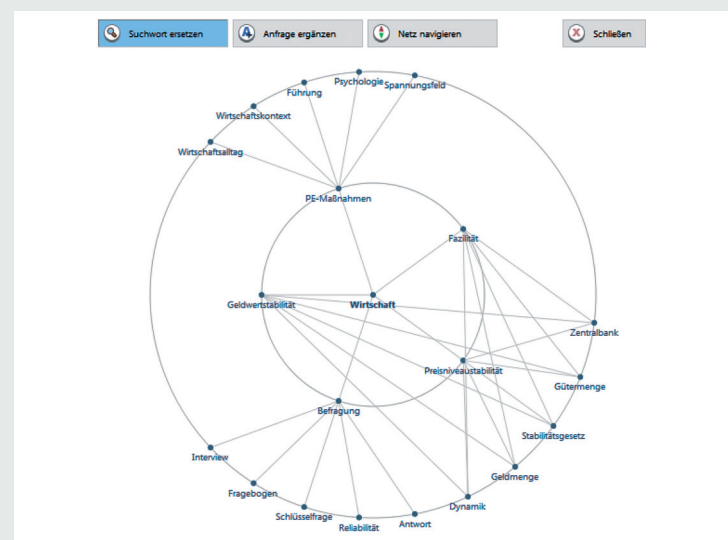


2. Synonymerweiterung

Abgesehen von der Anzeige von automatisch generierten Synonymen in den Alternativvorschlägen können auch Begriffe definiert werden, bei denen bereits die Suche selbst automatisch um Synonyme erweitert wird. Sucht ein Nutzer z.B. nach „Klempner“, so werden gleichzeitig zusätzlich „Sprengher“ und „Flaschner“ in die Suche aufgenommen. Damit vergrößert sich die Treffermenge.

3. Semantisches Netz

Bei dem semantischen Netz handelt es sich um eine Visualisierung der Beziehung zwischen den Begriffen im Index. Im Zentrum steht der Suchbegriff, im Ring darum finden sich verwandte Begriffe, für den zweiten Ring wird analog verfahren. Beispiel: „Wirtschaft“:



Dem Netz sind drei Funktionen hinterlegt:

Ein Klick auf ein Wort führt zu einer neuen Suchanfrage.

Die Suchanfrage kann auch um diesen Begriff erweitert werden.

Als dritte Funktion kann im semantischen Netz navigiert werden, d.h. der ausgewählte Begriff rückt in das Zentrum und die verwandten Begriffe und somit das Netz werden neu berechnet.

Screenshots: Thorsten Weise, Verlagsberatung www.weise-verlagsberatung.de