

PROZESSE & ORGANISATION

# Auf Sinnsuche

## Wie semantische Technologien den Arbeitsalltag erleichtern

Von Thomas Klein



Das Internet ist 2007 auf die astronomische Größe von 281 Milliarden Gigabyte angewachsen. Für den Nutzer wird es angesichts dieser Informationsflut immer schwieriger, Wichtiges von Unwichtigem zu unterscheiden. Unternehmen, die es schaffen, ihre Daten zu sortieren, verschaffen sich einen entscheidenden Vorsprung. Immer mehr Betriebe setzen daher bei ihrem Wissensmanagement auf semantische Technologien.

Im Jahr 280 vor Christus war das Weltwissen im Umfang von 100.000 Buchbänden in der Bibliothek von Alexandria zu finden. Heute – 2290 Jahre später – ist das Internet die neue Weltbibliothek. Es erweitert sich täglich im Umfang von 15 Milliarden 1.000-seitigen Büchern. Das Problem: Wie kann man diese Informationen so filtern, dass sie hilfreich sind? Heute gibt es eine Möglichkeit, die antiken Wissensdurstigen verwehrt war: Suchmaschinen. Auf Knopfdruck spu-

cken sie sekundenschnell Tausende von Treffern aus. Leider häufig die Falschen. Ein Beispiel: Man sucht für einen Artikel über semantische Technologien danach, wie groß die Gesamtmenge der Daten im Netz ist und wie viele Buchbände das umgerechnet ergibt. Gibt man die Suchbegriffe „Datenmenge“, „Internet“ und „Buchbände“ bei Google ein, erhält man als ersten Treffer eine Publikation zu den Beteiligungsverhältnissen in Deutschland. Der zweite enthält die Ausarbeitung eines literaturwissenschaftlichen Workshops. Was fängt man damit an? Wie kommt Google auf den Gedanken, dass der Suchende gerade hieran Interesse haben könnte? Das Problem liegt im Wesen des Suchalgorithmus der gängigen Suchmaschinen begründet: Sie suchen nach Häufigkeiten von Schlagwörtern und Verlinkungen. Über die Relevanz der gefundenen Dokumente gibt das nur begrenzt Auskunft. Synonyme Begriffe finden sie nicht, ebenso wenig berücksichtigen sie den Kontext der Suchanfrage und des Suchenden. Nicht selten sind die Treffer daher, wie in diesem Fall, unpassend und nicht zu verwenden.

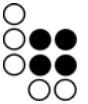
Mit ähnlichen Problemen sehen sich häufig die Mitarbeiter eines Unternehmens konfrontiert, wenn sie interne Informationen verarbeiten sollen. Denn: Was kümmert es den Vertriebler, welche Produktionsschritte das Ventil durchlaufen hat, bevor es beim Kunden ankommt? Der Ingenieur interessiert sich umgekehrt nicht für den Lieblingswein des Einkäufers seines Kun-

den. Das Problem also ist, dass Mitarbeiter zwar eine Menge Informationen bekommen – aber nicht unbedingt diejenigen, die sie eigentlich aktuell benötigen.

Eine Studie des Analystenhauses IDC ergab, dass sogenannte Wissensarbeiter, also solche, die vor allem auf intellektueller, weniger dafür auf körperlicher Ebene arbeiten, zwischen 15 und 35 Prozent ihrer Zeit mit der Suche nach Informationen verbringen und dabei in höchstens der Hälfte der Fälle erfolgreich sind. Mitarbeiter, die nicht über alle Fakten verfügen, treffen häufig keine oder die falschen Entscheidungen, was die Unternehmen teuer zu stehen kommen kann. Die Suche nach Informationen ist also kein Nebenkriegsschauplatz, sondern kann über Wohl und Wehe einer Firma entscheiden.

### Suche und Sinn verknüpfen

Betriebe müssen ihr firmeninternes Wissen daher derart speichern, strukturieren und aufbereiten, dass es den Mitarbeitern schnell und in hinreichender Qualität zur Verfügung steht. Das Ziel ist eine Suchtechnik, die den Sinn einer Anfrage versteht. In zunehmender Zahl setzen Unternehmen daher auf semantische Technologien. Der Ziegelproduzent Wienerberger beispielsweise organisiert sein Unternehmenswissen in einem sogenannten semantischen Netz. Diese Technologie verknüpft die Einzelinformationen derart, dass die Rechner eine Art Sinnzusammenhang erkennen. Die



Zettelwirtschaft und Datenschwungel waren gestern. Dank semantischer Technologien werden Informationen in Zukunft intelligent und individuell sortierbar.

Knotenpunkte eines solchen Netzes bilden die sogenannten Objekte: Ein Objekt kann die Firma A sein, der die Objektkategorie „Unternehmen“ zugeordnet ist. Darüber hinaus können einem Objekt auch Kategorien wie „Mitarbeiter“, „Projekte“, „Dokumente“, „Kompetenzbereiche“ und so weiter zugeordnet werden. Die Objekte werden unter Gesichtspunkten wie „verantwortlich für“, „im Auftrag von“, „hat Expertise in“ und Ähnlichem miteinander in Verbindung gesetzt.

Das System ordnet also nicht nur ein, um welche Form von Information es sich handelt, sondern auch, in welcher Beziehung die Objekte zueinander stehen.

Sobald bei Wienerberger ein neues Projekt startet, verschafft sich der Leiter mit Hilfe semantischer Technologien schnell einen Überblick über das vorhandene Vorwissen: Gab es vergleichbare Projekte in der Vergangenheit? Welche Schwierigkeiten tauchten auf? Welche Mitarbeiter wirkten dort mit, und wer bringt die nötigen Fähigkeiten für das Projekt mit? Bei der Schlagwortsuche vergleicht die Suchmaschine nicht nur Zeichenketten in der Anfrage mit den Zeichenketten in den durchsuchten Dokumenten, sondern sie „versteh“ auch den Sinn der Anfrage und bereitet die Informationen dementspre-

chend auf. Dabei bezieht sie auch den Kontext des Suchenden ein: Meldet sich der Ingenieur in seiner Funktion als Ingenieur an, erhält er bei seiner Suche eben keine Angaben über den bevorzugten Rotwein von Einkäufern.

Das ist tatsächlich so einfach, wie es klingt. Etwas komplizierter ist es dagegen, ein semantisches Netz erst einmal aufzubauen. Im ersten Schritt müssen die Daten strukturiert und zueinander in die passende Relation gesetzt

werden. Gemeinsam mit dem Hersteller der Software generieren die Fachanwender die Objekte des Netzwerks. Das erfordert einiges an manueller Arbeit, da die Objekte einzeln und von Hand definiert werden müssen. In den meisten Unternehmen steht allerdings bereits dank Kunden- und Projektdatenbanken, SAP-Systemen oder Excel-Dateien eine große Menge an gruppierten Informationen zur Verfügung: Aus diesen Datentöpfen bedienen sich die Programmierer, die die Schnittstellen zu den semantischen Netzen knüpfen.

Viele Unternehmen starten ihr internes Netz mit nicht mehr als hundert Objekten und Relationen. Ist das System einmal in Gebrauch, wächst und lernt es mit jeder Anwendung: Der einzelne Nutzer ist angehalten, das System über die Benutzeroberflä-

che zu pflegen. Er bewertet Suchergebnisse, knüpft neue Relationen und verlinkt weitere Dokumente. Neben der zentralen Pflege durch die Administratoren sind semantische Netze also in erster Linie kollaborativ.

Die nächste Revolution kommt bestimmt. Semantische Technologien werden das Web 3.0 aller Voraussicht nach dominieren. Doch obwohl IT-Giganten wie Google und Microsoft bereits mit neuen semantischen Suchmaschinen experimentieren, wird noch viel Zeit vergehen, bis das Internet mit Hilfe der Semantik eine neue Entwicklungsstufe erreicht. Dafür ist die zu strukturierende Datenmenge zu groß. Auf der internen Ebene von Unternehmen ist die Technologie aber heute schon nutzbar und trägt – wenn auch verbunden mit einigem Initialaufwand – dazu bei, dass Betriebe ihre Wissenssuche effizienter gestalten. ■

t.klein@faz-institut.de

### ESSENTIALS

- Die schnelle und erfolgreiche Suche nach Informationen im internen Netzwerk ist ein wichtiger strategischer Faktor für Unternehmen, den sie nicht unterschätzen sollten
- Semantische Technologien im Internet sind noch nicht ausgereift. In internen Netzwerken aber ist die Entwicklung weit vorangeschritten und kann Unternehmen einen Wettbewerbsvorteil verschaffen
- Eine semantische Suchmaschine hilft dabei, Wissen und Informationen je nach ihrer Relevanz zu filtern

Einmal im Gebrauch,  
wächst das System mit  
jeder Anwendung.