

Ontologiebasiertes Informationsmanagement zur integrierten Sicht auf Suchen, Dokumente und Wissen¹

Dr.-Ing. Wolf Engelbach, Dr.-Ing. Thomas Specht, Dr.-Ing. Anette Weisbecker

Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO)
Nobelstr. 12
70569 Stuttgart
wolf.engelbach@iao.fraunhofer.de

Abstract: Für die Informationsbeschaffung bei strategischen und operativen Projekten wird eine Softwarelösung² vorgestellt, die Internet-Suche, Dokumentenablage und Wissensspeicherung auf Basis von Ontologien miteinander verbindet.

Informationsbeschaffung für das Kundenmanagement

Viele strategische und operative Entscheidungen im Kundenmanagement fallen auf Basis zufällig erhaltener oder nicht systematisch ausgewerteter Informationen. Daraus können kostenträchtige Fehlentscheidungen hinsichtlich des Potentials von Märkten, der Position von Wettbewerbern oder der Qualität möglicher Partner erwachsen.

Die systematische Informationsaufbereitung ist daher eine zentrale Grundlage für dauerhaft erfolgreiches Kundenmanagement. Tatsächlich sind kontinuierlich vielfältige parallele Suchzusammenhänge für jedes Unternehmen relevant, beispielsweise über konkrete Kunden, Wettbewerbsprodukte, möglich ergänzende Dienstleistungen oder anvisierte neue Vertriebsmärkte. Diese Suchzusammenhänge werden im Folgenden als Projekte bezeichnet.

Bei der beabsichtigten Art von IT-unterstützter systematischer Informationsbeschaffung sind drei Sichten wichtig:

- Die Suche: welche Fragen muss ich mich welchen Suchwerkzeugen stellen, und wie merke ich mir erfolgreiche Kombinationen von Suchbegriffen?

¹ Mit freundlicher Genehmigung der Zeitschrift ERP Management (www.erp-management.de), in der dieser Beitrag in ähnlicher Form unter dem Titel „CRM-Unterstützung durch intelligentes Informationsmanagement: Internet-Suche, Dokumentenablage und Wissensspeicherung mit projektspezifischen Begrifflichkeiten“ zuerst veröffentlicht wurde.

² Die dargestellte Software-Lösung entstand im Projekt »AMI-SME« (Analysis of Marketing Information for Small and Medium Sized Enterprises; www.ami-sme.org, EU CRAFT Projekt 0005386). In dem EU-geförderten Technologietransferprojekt wirken 13 Partner aus Italien, Spanien, Polen, Österreich und Deutschland mit.

- Die Dokumente: wie speichere ich Dokumente so ab, dass ich sie für meine Auswertung zur Verfügung habe – und meine Notizen dazu ebenfalls?
- Das Wissen: Was habe ich aus der Sichtung der Suchergebnisse und Dokumente an Wissen gewonnen? Wie helfen mir diese Erkenntnisse bei der Beantwortung meiner Anliegen oder bei der Schärfung meiner Suchanfragen?

Für verschiedene Suchzusammenhänge, die mit Kundenmanagement in Beziehung stehen, sind immer wieder ähnliche Begriffe und Fragen relevant. Bei Geschäftsbeziehungen im B2B-Markt beziehen sich beispielsweise viele Informationsbedürfnisse auf Unternehmen:

- Zu einzelne Unternehmen sind ähnliche Informationen entscheidend, z.B. Mitarbeiterzahl und Firmensitz, des weiteren sind branchen- oder projektspezifisch andere Attribute von Interesse, z.B. die Vertriebsorganisation.
- Individuelle Kommentierungen und Bewertungen sollen es erlauben, zu einem Unternehmen Einschätzungen und offene Fragen festzuhalten.
- Relationen von einem Unternehmen zu anderen Informationstypen, z.B. Produkten, sollen einen Überblick zum erworbenen Wissen verschaffen: Unternehmen A vertreibt Produkt x, Unternehmen B setzt Produkt Y ein.
- Die Unternehmen sollen dabei nach bestimmten Untergruppen wie Wettbewerber, Partner oder Interessenten differenziert werden können, für die ggf. andere Attribute gepflegt werden.

Konzept für ein integriertes IT-System

Basierend auf den skizzierten Anforderungen wurde ein IT-System zur Unterstützung der Mitarbeiter in Unternehmen bei der Informationsbeschaffung in CRM-Prozessen gestaltet, das folgende Eigenschaften besitzt:

- eine intuitiv verständliche und einfach erweiterbare konsistente begriffliche Ordnungsstruktur mit unternehmens- bzw. projektspezifischer Differenzierungsmöglichkeit,
- eine über diese Begrifflichkeiten strukturierte Verbindung der Sichten von durchgeführten Suchen, gefundenen Dokumenten und extrahiertem Wissen sowie
- eine klare Trennung projektbezogener Ablageräume bei gleichzeitiger Integrationsmöglichkeit und redundanzfreier Speicherung von gefundenen Dokumenten.

Begriffswelt und Projektstruktur

Das IT-System ist zur Unterstützung der Internationalisierung von mittelständischen Unternehmen entstanden. Genau für diesen Kontext wurde daher eine systematische Begrifflichkeit („Ontologie“) erarbeitet, die zur Strukturierung der Suchen, Dokumentenablage und Wissensdatenbank dient.

Diese Ontologie ist als Datei im OWL-Format (Ontology Web Language) abgelegt. Mit jedem neuen Projekt wird diese Datei in eine Datenbank als projektspezifische Kopie geladen. Für diese Projektkopie werden alle Modifizierungen, Erweiterungen und konkreten Ergebnisse gespeichert. Dadurch steht die unveränderte Ausgangsdatei auch für folgende Projekte zur Verfügung.

Jedem Anwender steht es frei, auf Grundlage der bestehenden Ontologie zur Internationalisierung, auf Basis einer projektspezifischen Kopie, aus branchen- oder unternehmensspezifischen Begriffswelten oder ganz ohne Vorlage eine andere, für ihn passende Systemontologie mit verfügbaren Ontologiemodellierungswerkzeugen zu erstellen. Diese Ontologien sind inhaltlich auch für andere Schwerpunkte, beispielsweise Innovationsmanagement oder Produkteinführung denkbar und müssen nur bestimmte formale Anforderungen zur vollständigen Nutzung erfüllen.

Suche

Das System nutzt vorhandene Suchmaschinen und ist in diesem Sinne eine Meta-Suchmaschine; exemplarisch sind Google, Yahoo und A9 über Web Services integriert. In einer einfachen Suche müssen nur die Suchbegriffe eingegeben werden, zudem dient eine Namensgebung dem späteren Wiederauffinden und erlaubt ein Kommentarfeld eine Qualitätseinschätzung oder ähnliche Anmerkungen.

Neu in diesem System ist ein ontologiegestützter Suchassistent, durch den die in der projektspezifischen Begriffswelt enthaltenen Ausdrücke zur Spezifikation der Suche herangezogen werden können. Für einzelne Begriffe (z.B. „Forschungseinrichtung“) können zudem alle Unterbegriffe (z.B. „Private Forschungsvereinigungen“ und „Universitäten“) oder gleichzeitig alle konkreten Instanzen (z.B. „Fraunhofer IAO“ und „Universität Stuttgart“) für die Definition einer Suchanfrage übernommen werden.

Bei konkreten Instanzen lassen sich zudem gleichzeitig alle gepflegten Synonyme in eine neue Suche integrieren (z.B. „IAO“, „Fraunhofer-Institut für Arbeitswirtschaft und Organisation“). Auch mehrere solcher Sammel-Auswahlen lassen sich kombinieren, und eine manuelle Weiterbearbeitung ist jederzeit möglich.

Auf bereits ausgeführte Suchen und ihre Ergebnisse kann über eine projektspezifische Suchhistorie und eine Sessionhistorie zurückgegriffen werden. Besonders für wiederholt durchgeführten Suchen ist es hilfreich, dass sich auch nur die neuen Resultate anzeigen lassen.

Bild 1: Suchansicht mit Suchhistorie

Dokumente, die beim eigenständigen Surfen im Internet in Datenbanken gefunden werden, lassen sich durch Suchmaschinen nicht erfassen. Für das geschilderte System besteht daher eine Browser-Erweiterung, mit deren Hilfe eine Integration solcher Dokumente in ein konkretes Such-Projekt ebenfalls möglich wird. Insbesondere für Inhalte, die erst über Authentifizierung oder in Zusammenhang mit Zahlungssystemen zugänglich sind (beispielsweise Patentdatenbanken, oder Online-Journals), ist dies ein deutlicher Mehrwert gegenüber herkömmlichen Metasuchmaschinen.

Dokumentenablage

Die gefundenen Suchergebnisse lassen sich selektieren und weiter bearbeiten. Für jedes Ergebnis werden automatisch Metadaten extrahiert, die in einer Ansicht mit Dokumentendetails angezeigt werden können; weitere Angaben zu den Ergebnissen können manuell editiert werden, beispielsweise zum Autor. Eigene Kommentare oder Zitate lassen sich direkt zum Dokument speichern und erleichtern die spätere Auswertung. Das Dokument kann zur ergonomischen Bearbeitung auch in einem Pop-up gemeinsam mit den zentralen Metadaten- und Kommentarfeldern geöffnet werden.

Sobald die gefundenen Ergebnisse für das System herunter geladen werden, erfolgt automatisch eine tiefer gehende Analyse der Dokumenteninformationen, beispielsweise hinsichtlich der Größe, des Datei-Formates und der Sprache, sowie durch die Erzeugung eines automatischen Abstracts. Heruntergeladene Dokumente werden nur einmal systemweit gespeichert, selbst wenn sie in verschiedenen Projekten verwendet werden.

Die einzelnen Suchergebnissen – oder auch mehrere zugleich – können mit „Labels“ aus der projektspezifischen Ontologie versehen werden. Dadurch können zu einem späteren Zeitpunkt alle Dokumente, die für ein konkretes Unternehmen oder Land relevant sind, im Zusammenhang gesichtet werden.

Bild 2: Dokumentendetails mit Labeling (links) und Kommentierung (rechts)

Sind zu dem Thema eines Projektes bereits vor Nutzung dieses Systems relevante Dokumente gefunden worden, können diese über eine Up-load-Option aus dem eigenen Filesystem ins System integriert werden.

Wissensextraktion

Durch die Sichtung der Suchergebnisse und Dokumente extrahiert der Anwender persönliches Wissen, beispielsweise identifiziert er relevante Unternehmen auf dem Markt, oder er findet dessen Mitarbeiterzahl und Produktlinien heraus. Diese Informationen kann er im System ablegen und sie so für sich und andere Personen transparent und weiter nutzbar speichern.

Zu einzelnen Wissens-Typen wie Unternehmen kann der Nutzer darüber hinaus neue Attribute oder Untertypen definieren. Für konkrete Wissens-Einheiten wie Fraunhofer IAO werden dann die Werte für diese und die vordefinierten Attribute gespeichert sowie die Relationen zu anderen Wissens-Einheiten definiert. Diese Wissenssichten sind auch als Pop-up aufrufbar, um sie parallel mit der Ansicht der Suchergebnisse oder der Dokumente verwenden zu können.

Bild 3: Ontologiegestützte Wissensdatenbank mit Attributen und Synonymen

Architektur

Der Anwender greift mit einem Internet- Browser auf das System zu. Die Ontologie sowie die Benutzungsoberfläche sind mehrsprachig ausgelegt; deutsch, englisch, spanisch, italienisch und polnisch sind mittelfristig vorgesehen. Das System läuft in einem Tomcat Server und kann beliebige Datenbanken unterstützen.

Die Implementierung erfolgte mit Java unter Verwendung eines Frameworks mit zahlreichen Standardfunktionalitäten (Objectledge). Die Lösung ist multiuserfähig. Daher ist sowohl eine stand-alone-Variante möglich, bei der Client und Server auf dem selben Computer laufen (z.B. dem Laptop eines Consultants), als auch eine unternehmensinterne Client-Server-Anwendung und konzeptionell zudem ein Application Service Providing durch einen zentralen Anbieter oder mehrere Branchenverband oder Beratungsunternehmen.