

ConWeaver – Die semantische Suchmaschine

Um Unternehmenswissen verfügbar und auffindbar zu machen, genügt es nicht, Datenbanken und Dokumentenserver über Intranets und Portale nur an eine gemeinsame Oberfläche zu koppeln. Ohne logische Integration der Daten- und Dokumentbestände und eine intelligente Suche ist die gewünschte Information nur mit kenntnisreichen Frageformulierungen für jede einzelne Datenquelle und aufwändigen Strategien für den Abgleich der Suchergebnisse zu finden. Dieser Umstand wird von immer mehr Unternehmen erkannt. Hier setzt die semantische Suchtechnologie von ConWeaver an.

ConWeaver – Die Lösung für die semantische Unternehmenssuche.

ConWeaver ermöglicht eine intelligente Suche im Unternehmensportal oder Intranet durch die Integration von Daten- und Dokumentbestände mit Hilfe eines Wissensnetzes, auch semantisches Netz oder Ontologie genannt. Dieses Wissensnetz nutzt ConWeaver als logischen Suchindex für alle angeschlossenen Datenquellen. Der Nutzer adressiert mit nur einer Frageformulierung alle Daten- und Dokumentbestände und erhält eine vollständige, aggregierte und strukturierte Antwort.

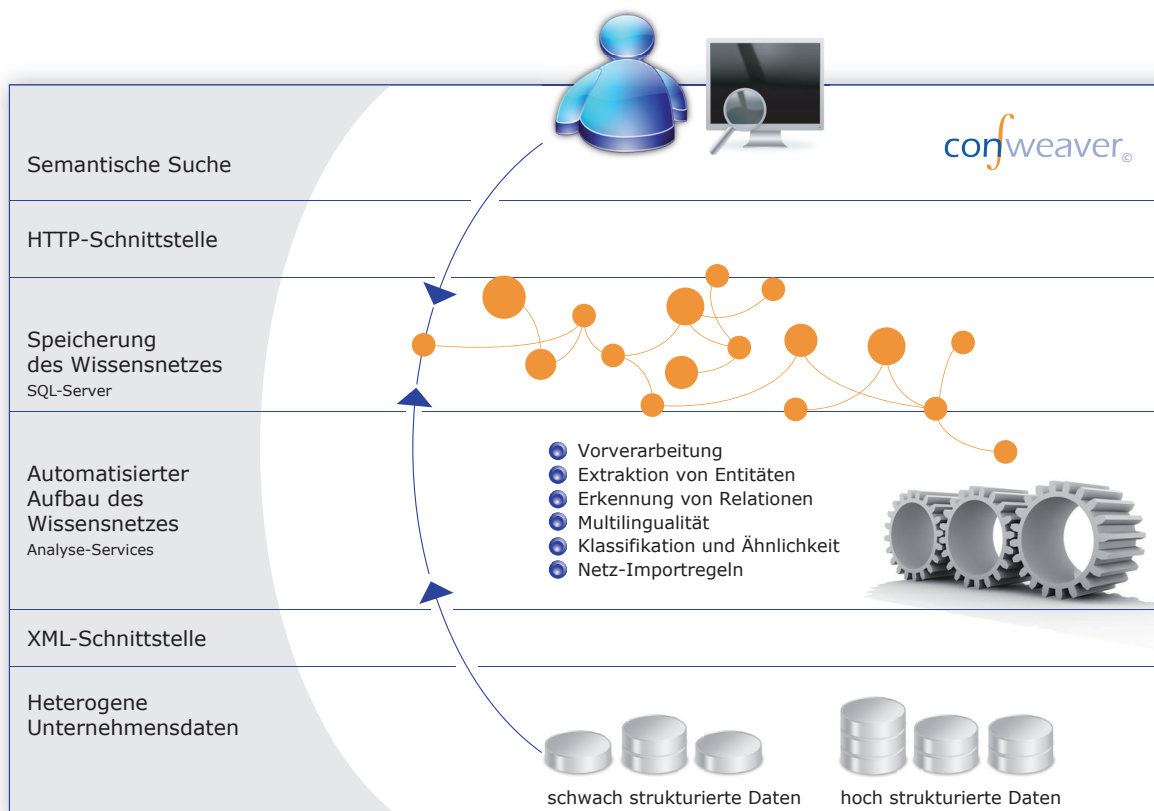
ConWeaver transformiert Daten in Wissen und unterstützt damit den semantischen Informationsprozess in Ihrem Unternehmen.

ConWeaver umfasst Analyse-Services für den automatisierten Aufbau von Wissensnetzten und Komponenten für die semantische Suche. Darüber hinaus ermöglicht die ConWeaver-Technologie ein Rapid Prototyping.

Automatisierter Aufbau von Wissensnetzen

Um Daten und Dokumente inhaltlich zu erschließen und das gewonnene Wissen zu vernetzen und zu integrieren, müssen unterschiedliche Datenmodelle, strukturierte Faktendaten und unstrukturierter Textinhalt aus den Quellsystemen zunächst auf ein einheitliches Schema abgebildet werden. ConWeaver setzt dabei Wissensnetze als Zuordnungsschemata ein. Diese Wissensnetze werden - und das ist die innovative Leistung von ConWeaver - mit Verfahren der künstlichen Intelligenz und des maschinellen Lernens automatisiert aufgebaut und automatisch gepflegt.

ConWeaver führt alle Daten und Dokumente Ihres Unternehmens ohne System- und Medienbrüche in einem automatisiert aufgebauten Wissensnetz zusammen und erschließt ihren Inhalt für eine hochqualitative unternehmensspezifische Suche.



Wissensnetze

Wissensnetze sind formalisierte Schemata, mit denen Wissen zu Geschäftsobjekten wie Kunden, Lieferanten, Produkten oder Projekten beschrieben werden kann. Ein Wissensnetz besteht aus Begriffen (Konzepten), die über bedeutungstragende Relationen miteinander verbunden sind. Sie bilden die sogenannte Schemaebene, also das Modell für die auf der Individuenebene angeordneten konkreten Individuen (Entitäten) und die sie verbindenden Relationen. Der manuelle Aufbau und die Pflege von Wissensnetzen gestalten sich allerdings sehr wissens-, zeit- und damit kostenintensiv.

Deshalb setzt ConWeaver einerseits innovative Technologien der statistischen und linguistischen Informationsextraktion und Kategorisierung ein, wenn die Datenquellen aus Texten bestehen. Andererseits kommen Best-Practice-Verfahren aus der Datenintegration zum Zuge, wenn es sich um strukturierte Daten handelt, wie sie in relationalen Datenbanken vorliegen. Hinzu kommen Methoden der maschinellen Übersetzung für den Aufbau und die Pflege multilingualer Wissensnetze und Methoden für die automatische Terminologieextraktion aus Internet-Quellen. Dieses gesamte Aufgabenspektrum obliegt den Analyse-Services, der semantischen Middleware von ConWeaver.

Automatische Informationsextraktion und Indexierung

Mit linguistischen und statistischen Text-Mining-Verfahren gewinnt ConWeaver Informationen, beispielsweise zu den Konzepten Person, Organisation, Projekt oder Thema aus Textdokumenten unterschiedlicher Formate. Mit Methoden der Kategorisierung und Klassifikation ordnet ConWeaver Dokumente vorgegebenen Themenkategorien zu oder organisiert sie in neuen Kategorieklassen. Verfahren der automatischen Übersetzung lösen das Mehrsprachenproblem und die Einbindung von Fachterminologien gewährleistet einen fachsprachlichen Zugang.

Linguistische und statistische Verfahren

Mit Hilfe weiterer linguistischer Analysemethoden wird der Wortschatz aus den Text-Mining-, Kategorisierungs- und Übersetzungsschritten weiter ausgewertet. Neuartige Verfahren ermitteln logische Beziehungen wie Ober- und Unterbegriffsrelationen zwischen Begriffen und erkennen, ob es sich um Begriffe, wie z.B. Person oder konkrete Individuen, wie z.B. die Person Hugo Schmidt handelt. Aufgrund dieser Charakteristika kann ein Begriff in einem semantischen Netz positioniert und automatisch in das Wissensnetz eingefügt werden.

Semantische Suche

Bei der semantischen Suche nutzt ConWeaver das Wissensnetz als integrierten Index, um in verteilten Datenbeständen zu suchen. Die fach-, umgangs- und fremdsprachlichen Verknüpfungen von Begriffen im Wissensnetz ermöglichen eine bedeutungsspezifische, mehrsprachige und formulierungsflexible Suche. Auf Grund der in Begriffen und Beziehungen hinterlegten Bedeutungen einerseits und der logischen Formalisierung andererseits sind schlussfolgernde und sprachverstehende Suchen auf den konkreten Individuen im Wissensnetz möglich. So können relevante Informationen selbst dann gefunden werden, wenn sie implizit in Dokumenten und Dateneinträgen vorhanden sind.

Die „geführte Suche“ ermöglicht durch Vorschläge zur kontextspezifische Einschränkungen oder Ausweitungen der Suche, insbesondere die Erschließung unstrukturierter Datenquellen.

Selbstverständlich schließt die intelligente ConWeaver-Suche auch alle Leistungsmerkmale einer linguistisch aufbereiteten Volltextsuche wie Trunkierung, Grundformreduktion, Schreibfehlertoleranz oder logische Operatoren ein.

Rapid Prototyping

Das graphische Interface und der modulare Charakter von ConWeaver zur Definition von Workflows für den automatisierten Aufbau von Wissensnetzen ermöglichen eine schnelle Umsetzung einer prototypischen Suchanwendung im Unternehmen. Sie erfolgt auf Basis eines zuvor mit dem Kunden definierten exemplarischen Datensatzes. Dieser so entstandene Prototyp vermittelt einen ersten Eindruck von den umfassenden Möglichkeiten einer mit ConWeaver aufgebauten semantischen Unternehmenssuche und kann als Kern einer sich daran anschließenden Projektrealisierung genutzt werden.

Mit ConWeaver werden strukturierte Faktendaten aus Datenbanken ebenso eingebunden wie Informationen aus unstrukturierten Texten, quantitative Daten ebenso wie qualitative.

ConWeaver kombiniert innovative sprachverstehende und schlussfolgernde Verfahren aus der künstlichen Intelligenz und dem maschinellen Lernen mit Best-Practice-Standards aus der Datenintegration und dem Information Retrieval.

Mit ConWeaver sind alle Datenquellen Ihres Unternehmens über eine einheitliche Suchoberfläche und Ergebnispräsentation zugänglich.

ConWeaver ermöglicht die schnelle Umsetzung prototypischer Suchanwendungen für Ihr Unternehmen.

ConWeaver
Fraunhoferstraße 5
64283 Darmstadt

Dr. Thomas Kamps
Fon +49 (0) 6151/155-651
Fax +49 (0) 6151/155-139
thomas.kamps@conweaver.de
www.conweaver.de