

Enterprise Search und Wissensmanagement

Kognitive Suche: Durchbruch für Enterprise Search

Mit Beginn der Ära des Cognitive Computing entstehen derzeit neue Lösungen, die leistungsstarke Indizierungstechnologie mit fortgeschrittenem Natural Language Processing (NLP) und Machine-Learning-Algorithmen kombinieren. So bildet sich eine neue Generation des Enterprise Search heraus.



Zu den Vertretern dieser Richtung gehört der Hersteller Sinequa. Mit Hans-Josef Jeanrond, Chief Marketing Officer bei Sinequa, sprach »manage it« über das Konzept der kognitiven Suche.

Gescheiterte Projekte, zu hohe Erwartungen – Enterprise Search und Wissensmanagement scheinen etwas aus der Publikumsgunst gefallen zu sein. Kognitive Suche und Machine Learning sollen ihnen jetzt zu neuem Glanz verhelfen. Wie kann das funktionieren?

Bei der kognitiven Suche verbindet sich Indexierung mit der Verarbeitung natürlicher Sprache (NLP) und Maschine Learning. Diese Kombination erlaubt eine viel tiefere Analyse von Inhalten und Nutzerverhalten als einfache Key-

mit rund 80 Prozent der Großteil geschäftlicher Daten in unstrukturierter Form vor, als E-Mail, PDF oder in Archivsystemen. Solche Quellen enthalten eine Fülle von Informationen, die nicht kodifiziert sind und sich nicht in bloßen Zahlen ausdrücken lassen. Mit kognitiver Suche und semantischer Inhaltsanalyse gelingt es, in ihnen verborgene Korrelationen zu ermitteln und semantisch ähnliche Dokumente aufzuspüren. Es werden auch solche Inhalte erfasst, die einen gesuchten Begriff unter Umständen gar nicht enthalten, aber ein verwandtes Thema behandeln und deshalb interessant sein könnten.

Moderne Systeme für die Enterprise-Search- und Big-Data-Analyse müssen daher beide Datenwelten, struk-

Cloud-Konnektoren zur Indizierung und Analyse von Daten aus AWS und Microsoft Office 365 mit Sharepoint Online, Exchange Online, Microsoft Onedrive und Yammer und so weiter.

Wie funktioniert das Einsammeln und Bereitstellen der relevanten Suchergebnisse bei Sinequa?

Die Software analysiert die Quellen mittels semantischer Methoden und NLP und indexiert ihre Inhalte. Der Index ergibt ein Datenreservoir, ein sogenanntes Logical Data Warehouse (LDW), das kontinuierlich durch die Ergebnisse der Analyse von Machine-Learning-Algorithmen angereichert wird. Zu den gebräuchlichsten Machine-Learning-Algorithmen in diesem



» **Kognitiv sind die Lösungen**, weil sie mit dem Anwender in einer natürlichen Art und Weise interagieren, weil sie selbstständig dazulernen durch Sammeln und Analysieren von Daten und indem sie proaktiv Verbindungen zwischen verwandten Daten aus verschiedenen Quellen herstellen. «

Hans-Josef Jeanrond, Chief Marketing Officer bei Sinequa

word-Suche. Eine Enterprise-Search-Software, die auf solche Techniken setzt, kann dem Nutzer also Informationen mit ständig verbesserter Relevanz und tiefere Einblicke innerhalb seiner täglichen Arbeitsumgebung zur Verfügung stellen. Kognitiv sind die Lösungen, weil sie mit dem Anwender in einer natürlichen Art und Weise interagieren, weil sie selbstständig dazulernen durch Sammeln und Analysieren von Daten – auch über das Nutzerverhalten – und indem sie proaktiv Verbindungen zwischen verwandten Daten aus verschiedenen Quellen herstellen, sowohl intern als extern.

Welche Art von Quellen meinen Sie hier?

Strukturierte und unstrukturierte gleichermaßen. In der Vergangenheit wurde viel in Business Warehouse und Business Intelligence investiert, um strukturierte Daten zu erschließen. Dabei liegt

turierte und unstrukturierte, zusammenführen, und dies gelingt eben durch kognitive Suche und Analyse. Es muss also eine Menge Intelligenz und Leistung unter der Haube eines Systems vereint werden, um relevantes Wissen oder Einsichten bereitzustellen.

Erfordert es nicht hohen Programmieraufwand, diese verschiedenen Datenquellen an die Enterprise-Search-Software anzuschließen?

Im Prinzip schon. Deswegen haben wir eine Konnektortechnologie entwickelt und liefern zum heutigen Stand 150 intelligente Konnektoren zu allen gängigen strukturierten und unstrukturierten Datenquellen am Markt mit aus. Wir nennen sie »intelligent«, nicht weil sie Künstliche Intelligenz benutzen, sondern weil sie die Struktur von Quelldaten kennen. Dazu gehören auch Anschlüsse für MapR Hadoop,

Kontext gehören Collaborative Filtering und Empfehlungen, Klassifizierung über Beispiele, Clusterbildung, Ähnlichkeitsberechnungen für unstrukturierte Inhalte und prädiktive Analyse. Sie schaffen einen Mehrwert durch kontinuierliche Verfeinerung und Verbesserung der Suchergebnisse.

Auf dieser Grundlage lassen sich dann sogenannte Search Based Applications oder Info-Apps zur dezidierten Informationsbeschaffung für einzelne Nutzer(gruppen) erstellen. Das LDW schützt Entwickler von Info-Apps vor den Komplexitäten der ursprünglichen Datenquellen und bietet einen einheitlichen Datenzugriff. Sie können so viel schneller und effizienter agieren.

Können Sie ein paar Beispiele solcher Info-Apps geben?

Sicher. Der Pharmakonzern Astra-Zeneca hat seit Einführung unserer

Kognitive Suche

Quelle: Sinequa

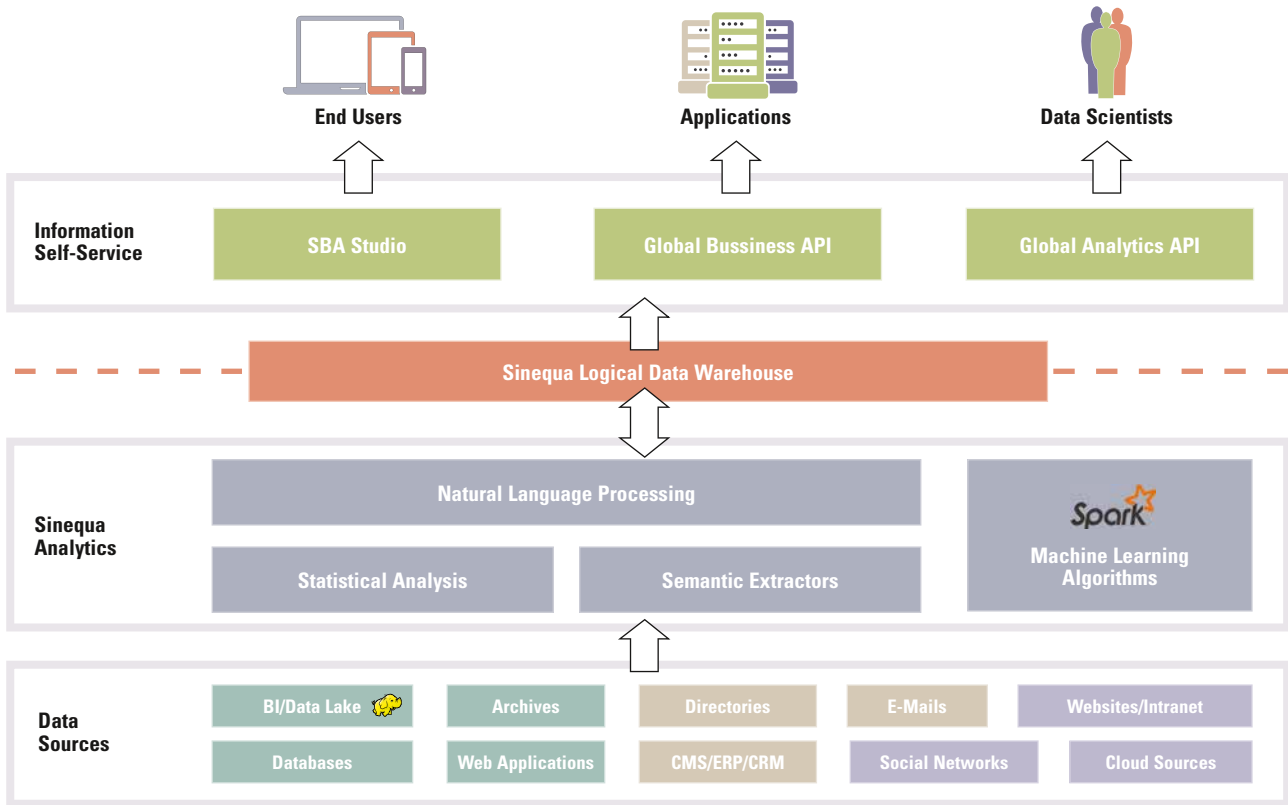


Abbildung 1: Die Architektur der Sinequa-Suche.

» **Marktanalysen zeigen,** dass datengetriebene Organisationen allein durch fehlende Informationen viele Millionen Euro pro Jahr verlieren. «

Analyseplattform in seiner F&E-Abteilung einen regelrechten »AppStore« für solch individualisierte Informationsbeschaffungen aufgebaut. Die App »R&D Pulse« etwa umfasst rund 110 Millionen Dokumente aus externen Datenquellen wie Publikationen, Inhalten branchenspezifischer Informationsquellen wie Embase, Medline, Scopus, klinischen Testberichten usw., sowie internen Datenquellen wie beispielsweise Sharepoint, Documentum.

Aus all diesen Quellen liefert R&D Pulse neueste wissenschaftliche Informationen aus dem Fachgebiet eines Anwenders, so dass er einfach auf dem Laufenden bleiben kann. Über »R&D Experts« lassen sich Expertennetzwerke

zu einem dezidierten Thema zusammenstellen. Dies fördert die Zusammenarbeit und vermeidet redundante Projekte. »R&D Intelligence« identifiziert neue Methoden zur Zielidentifikation und Neupositionierung von Medikamenten.

Was sind weitere praktische Anwendungsfälle?

Ein Kunde aus der Pharmabranche stellt mit unserer Cognivite-Search-Plattform zum Beispiel Expertenteams zu bestimmten Themen zusammen. Andere kreieren 360-Grad-Ansichten zu Produkten, Zulieferketten, Kunden und Wettbewerbern oder sie analysieren die Navigation der Besucher auf

ihrer Website und die von diesen initiierten Suchanfragen.

Zu guter Letzt: Warum sollte ein Unternehmen in kognitive Suche und Analyse investieren?

Marktanalysen zeigen, dass datengetriebene Organisationen allein durch fehlende Informationen viele Millionen Euro pro Jahr verlieren. Indem sie heterogene Datenmengen aus verschiedensten Quellen in verwertbares Wissen verwandeln, sparen sie also bares Geld und stärken ihre Wettbewerbsfähigkeit. Diese Erfahrung haben alle unsere Kunden gemacht, angefangen von Airbus über AstraZeneca, Atos und Crédit Agricole bis hin zu Siemens. »