

Hans-Josef Jeanrond

## Hat Enterprise Search noch eine Zukunft?

**„Enterprise Search“ ist in Verruf gekommen: Zu viele hochgesteckte Erwartungen, die nicht erfüllt wurden, zu viel Frustration bei den Anwendern über irrelevante Suchergebnisse. Daher suchen Software-Firmen und Marktanalysten neue Namen für den Markt der zukünftigen Systeme, die Enterprise Search ablösen sollen. Gartner hat sich für „Insight Engines“ entschieden und damit „Search“ völlig verschwinden lassen.**



© Rainer Sturm/PIXELLO – pixello.de

Trotzdem berichten die Analysten von Gartner, Forrester und IDC, dass sie noch immer viele Anfragen von ihren Kunden zum Thema Search bekommen. Mike Gualtieri von Forrester will daher auch weiter an dem Begriff festhalten. Um gleichzeitig darauf hinzuweisen, dass Search sich weiter entwickelt hat, wird er wohl künftig von „Cognitive Search“ sprechen.

Insight Engines oder kognitive Such- und Analyse-Plattformen, was werden Software-Systeme mit diesen Etiketten liefern? Wo liegen Anwendungsfelder mit dem höchsten ROI? Und was sind mögliche Risiken für Unternehmen?

### → Insight Engines oder kognitive Suchplattformen?

Das folgende Bonmot von Kevin Kelly, dem Gründer und Redakteur des Wired Magazins, kann als Überschrift für den Übergang von Enterprise Search zu Insight Engines oder kognitiven Suchplattformen dienen: „The business plans of the next 10,000 startups: Take X and add AI.“ Das hieße in unserem Fall: Man nehme Enterprise Search und füge künstliche Intelligenz (KI) hinzu.

Ganz so einfach ist es natürlich nicht. Man kann nicht einfach KI auf ein veraltetes System „aufstäuben“, um ihm eine neue Zukunft zu geben. Eine Plattform für kognitive Suche und Analyse

braucht Kreativität und Innovation, eine solide Architektur und eine skalierbare Implementierung. Die Formel für erfolgreiche Suche könnte lauten: Cognitive Search = Search + NLP + KI/ML. Doch was bedeutet das konkret?

### → Verarbeitung natürlicher Sprachen

Mancherorts wird erklärt, dass die Fähigkeit zur Verarbeitung von Texten (Natural Language Processing, kurz: NLP) bereits in den alten Enterprise-Search-Systemen enthalten war und in den neuen Lösungen von KI-Algorithmen ohne Rückgriff auf die klassische Linguistik geliefert wird. Beides trifft jedoch nicht ganz zu: Die meisten der alten Systeme konnten nur rudimentäre Sprachanalyse liefern (und das nur in wenigen Sprachen). Sie konnten Schlüsselwörter aus vordefinierten Listen („White Lists“) erkennen und einfache linguistische Analyse betreiben. KI-Algorithmen zur Sprachanalyse können im weltweiten Web mit seinen gewaltigen Textsammlungen zu brauchbaren Ergebnissen führen, aber selten in Unternehmen oder Organisationen. Die statistische Basis von Unternehmensdaten ist schwächer, KI-Algorithmen liefern weniger gute Resultate.

Es sollte selbstverständlich sein, dass „natürliche Sprachen“ sich nicht auf Deutsch, Englisch und „noch eine Spra-

che“ beschränken. Chinesisch, Japanisch und Koreanisch gehören zum Kanon der wirtschaftlich wichtigen Sprachen, ebenso wie Arabisch und eine Reihe von europäischen Sprachen. 15 bis 20 Sprachen sollte ein System, das auf den weltweiten Markt zielt, schon umfassen.

### → Künstliche Intelligenz und Machine Learning

Noch steht nicht fest, welche Machine-Learning (ML) -Algorithmen und welche Neuronale Netze optimale Resultate für die kognitive Suche erbringen und gleichzeitig den besten ROI in Unternehmen liefern. Die meisten Unternehmen experimentieren noch in diesem Bereich.

Ein Trend zeichnet sich aber bereits deutlich ab: Neuronale Netze werden mit gutem Erfolg verwandt

- in der Behandlung von gesprochener Sprache (Speech-to-Text),
- in der Erkennung von handschriftlichen Texten,
- in der automatischen Übersetzung,
- in der Bild- und Video-Analyse.

Die großen Akteure der IT-Branche – Amazon, Google, IBM und Microsoft – bieten bereits entsprechende Dienstleistungen in der Cloud an: Amazon Machine Images (AMI), Microsoft Azure Media Services, Google Vision, Google Natural Language, IBM Watson Developer Cloud mit Visual Recognition, Alchemy Language, Speech-to-Text.

So wird sich heute kaum noch ein Unternehmen für diese Funktionalitäten eigene Neuronale Netze konstruieren lassen. Die Stärke der großen Anbieter liegt gerade im enormen Volumen der ihnen zur Verfügung stehenden „Testdaten“ (Bilder, Audio- und Videofiles aus dem Internet) und in der bereits absolvierten Trainingsphase ihrer Neuronalen Netze. Kognitive Suchplattformen tun allerdings gut daran, diese Dienste nahtlos zu integrieren, so dass entsprechende Inhalte analysiert und indexiert werden können, ohne dass Anwender mit Systembrüchen konfrontiert werden.

Zu den Machine-Learning-Algorithmen, die zur kognitiven Suche auf Anhub nützlich erscheinen, gehören z. B.

- automatische Clusterbildung (Einordnung von Inhalten in Kategorien),
- automatische Klassifizierung anhand von Beispielen ohne definierte Regeln (Classification by Example),
- Empfehlungen („wenn Sie das interessiert, sollte Sie auch jenes interessieren“),
- prädiktive Algorithmen (Predictive Analytics, Regression) und
- Ähnlichkeitsberechnungen (streng genommen kein ML-Algorithmus, aber durch den hohen Rechenaufwand auf heuristische Versionen beschränkt und zu den ML-Algorithmen geschlagen).

Die meisten dieser Algorithmen sind seit Langem bekannt, konnten aber mangels Performance der Computersysteme nicht eingesetzt werden und wurden Anfang der 80er Jahre auf Eis gelegt. Da nach Moore's Law heutige Computer jedoch gut 16 Millionen mal leistungsfähiger sind als ihre Vorgänger vor 36 Jahren, ist Machine Learning wieder zu einem „heißen“ Thema geworden.

ML-Algorithmen werden heute im Allgemeinen in SPARK-Umgebungen programmiert. Kognitive Suchplattformen sollten daher solche Umgebungen integrieren, gleichzeitig aber Schnittstellen anbieten, so dass Unternehmen ihre eigenen SPARK-Cluster einsetzen können. Das erlaubt ihnen auch den Einsatz von ML-Algorithmen, die in einer Suchplattform nicht enthalten sind – ein interessanter Freiheitsgrad, solange sich noch kein „Standardsatz“ von ML-Algorithmen für Suchplattformen herauskristallisiert hat.

### → Die Kombination macht den Unterschied!

NLP, KI und „klassische“ Suche müssen in eine innovative und kohärente Architektur eingebettet werden, um die bestmöglichen Resultate für Unternehmen zu liefern. Sie dürfen nicht einzeln oder nebeneinander eingesetzt, sondern müssen verzahnt werden. So liefern z. B. ML-Algorithmen wesentlich schneller wesentlich bessere Ergebnisse, wenn sie auf einem „logischen Datawarehouse“ aufsetzen können, in dem schon Begriffe und Beziehungen zwischen Begriffen aus den Rohdaten und Texten extrahiert worden sind.

### → Anwendungsszenarien

Einige Anwendungen kognitiver Suche sind bereits „klassisch“, auch wenn das Gebiet noch relativ neu ist. Dazu zählen die 360-Grad-Sicht, zum Beispiel auf einen Kunden, und das Aufzeigen von Experten-Netzwerken zu einem bestimmten Thema. In beiden Fällen müssen strukturierte und unstrukturierte Daten zusammen analysiert werden, die NLP-Komponente muss Fachvokabular „verstehen“ können und Beziehungen zwischen (Fach-) Begriffen herstellen. Bei der automatischen Klassifizierung des Posteingangs wird neben NLP zusätzlich Machine Learning eingesetzt: Klassifizierung anhand einer Beispielmengende. Auch Empfehlungen á la „Leute, die abc gekauft haben, haben auch xyz gekauft“ gehören zu den klassischen Anwendungsbereichen. Diese Empfehlungen können sehr banale Auswertungen von „Einkaufskörben“ sein, aber auch anspruchsvoller, mit der Berechnung von „Ähnlichkeiten“.

Darüber hinaus gibt es auch bereits wesentlich innovativere Einsatzmöglichkeiten. Dazu zählen:

- **Einhaltung von Compliance-Richtlinien**  
In stark regulierten Branchen, wie z. B. der Pharmaindustrie, müssen Unternehmen die Publikation neuer Regelungen in vielen Ländern und Sprachräumen überwachen, feststellen, ob sie für die eigenen Produkte relevant sind, und wenn ja, die Informationen an die kompetenten Stellen im Unternehmen weiterleiten. Kognitive Suchsysteme analysieren, worum es bei neuen Regeln geht, und nutzen die in der Vergangenheit von menschlichen Experten kategorisierten Dokumente, um mithilfe von Machine Learning (Classification by Example) neue Dokumente an die richtigen Personen im Unternehmen weiterzuleiten.
- **Auswertung von Patienten-Video-Interviews bei klinischen Tests**  
Interviews zwischen Ärzten und Probanden bei klinischen Tests erfolgen oft per Video-Interview. Die Ergebnisse werden in dedizierten Formularen festgehalten. Die Unterhaltungen enthalten aber oft wesentlich mehr Informationen als in den Formularen erfasst wird. Eine Übertragung der sprachlichen Inhalte in

Texte (Speech-to-Text) macht diese Inhalte kognitiven Suchsystemen zugänglich, so dass wertvolle zusätzliche Informationen ausgewertet werden können.

- **„Voice of the Customer“**  
Auswerten sozialer Netzwerke, Herausfiltern positiv oder negativ besetzter Themen.
- **Auswertung von Messdaten und Fehler-Reports.**  
Komplexe Systeme, wie Flugzeuge oder Kraftwerke, erzeugen einen ständig wachsenden Strom von Messdaten aller möglichen Komponenten. Hier können kognitive Such- und Analysensysteme ihre Stärken der kombinierten Analyse besonders gut ausspielen.

### → Fazit

Die Anwendungsfälle für kognitive Such-Systeme oder Insight Engines sind unglaublich vielfältig. Es gibt noch kaum Best Practices, die einfach befolgt werden können. Unternehmen müssen selbst kreativ sein, wenn sie sich im neuen Umfeld des Cognitive Computings und speziell mit kognitiven Suchsystemen einen Wettbewerbsvorteil verschaffen wollen.

### → Der Autor



Quelle: Kevin Swift

*Hans-Josef Jeanrond ist Chief Marketing Officer (CMO) von Sinequa. Er verfügt über langjährige Erfahrung im IT-Marketing. Als ausgebildeter Informatiker (Studium in Saarbrücken, Oxford und Edinburgh) sieht er sich als Vermittler zwischen der technischen IT-Welt und der Welt der Anwender, die am Wert von IT-Innovationen für ihr Unternehmen interessiert sind.*

✉ jeanrond@  
wissensmanagement.net