

INNOVATIONSPREIS 2010

Sieger: Lufthansa Technik AG und TextTech GmbH

Insight Mining

Modernes Data Management auf der Grundlage der Text Mining-Technologie

Wie geht man effizient mit der Vielzahl an Informationsquellen um, mit denen Betriebsmarktforscher jeden Tag konfrontiert werden? Wie kann man Zugriff zu Informationen bekommen, wenn diese gebraucht werden?

Vor dem Hintergrund der Schwierigkeiten im Umgang mit Informationsquellen aller Art skizzierte das Marktforschungsteam der Lufthansa Technik AG (LHT) im Dezember 2008 folgende Idee:

- Ein Datenpool, der mit Dokumenten gefüllt wird (Zeitschriften, Newsletter, Analystenreports, Konferenzpräsentationen, etc.), der bei Bedarf nach Informationen effizient durchsucht werden kann.
- Es sollte möglich sein, Beziehungen zwischen Begriffen zu finden, die vorher nicht bekannt waren, aber auch vermutete Beziehungen zu überprüfen.
- Darüber hinaus sollte eine Archivfunktion bestehen, damit man auf die vollständigen Originaldokumente zurückgreifen kann.

Ein EDV-Tool mit den oben beschriebenen Eigenschaften gab es bisher nicht. Es blieb also nur eines übrig: eine neue Anwendung zu entwickeln.

Zusammen mit TextTech entwickelte die LHT ab Januar 2009 erste Lösungsansätze und Ideen. Im Frühjahr 2009 konnte ein erster Prototyp präsentiert werden. Die ersten Tests fielen ausgesprochen positiv aus, ab August 2009 wurde mit der Realisierung des Tools Insight Mining begonnen.

Das Neue an diesen für den BVM preisverdächtigen Ablagesystemen ist die Kombination mit „Text Mining“, das mit statistischen und linguistischen Methoden versteckte Beziehungen zwischen Texten aufzeigen kann.

Es handelt sich bei diesem Instrument um eine Entwicklung der TextTech, die LHT war Launching Customer und begleitete die Entwicklung mit Input aus der Praxis der betrieblichen Marktforschung. Dadurch wurde es möglich, ein Werkzeug zu entwickeln, welches den hohen Anforderungen der betrieblichen Marktforschung gerecht wird. Insight Mining sollte nicht nur die Bedürfnisse der LHT abdecken, sondern von Anfang an eine Implementierung bei anderen Unternehmen erlauben.

Ein reines Ablagesystem zu schaffen, stellt keine besondere Herausforderung dar und ist auch nicht mehr zeitgemäß. Das Tool ist innovativ, da erstmals Text Mining-Methoden für ein Knowledge-Management-System verwendet werden. Bisher wurden diese Methoden in erster Linie für Web 2.0-Analysen verwendet.

Text Mining ist der Prozess der Zusammenstellung, Organisation und Analyse großer textueller Dokumentensammlungen mit statistischen und linguistischen Mitteln sowie der Entdeckung versteckter Beziehungen zwischen Texten und Textfragmenten mit dem Ziel der Wissenserweiterung. Somit besteht eine methodische Verwandtschaft mit dem Data Mining.

Das Analysetool von Insight Mining ist das Herzstück der Anwendung. Durch eine graphische Benutzeroberfläche ist ein schnelles und intuitives Surfen durch den Datenpool möglich. Die Suchoberfläche wird durch eine Vielzahl an Suchoptionen ergänzt, die eine effiziente Recherche erlauben. Darüber hinaus erfüllt das Instrument durch seinen modularen Aufbau noch weitere Anforderungen aus der Praxis, unter anderem durch die Möglichkeit, user generated content einzubinden.

Die Entwicklung von Insight Mining war in vielerlei Hinsicht eine Herausforderung für alle Beteiligten, unter anderem durch die Entwicklung neuer, praxisgerechter Berechnungsalgorithmen. Am Ende gelang es, ein Tool zu entwickeln, welches Suchanfragen um den Faktor vier bis sechs beschleunigt und daher ausgesprochen wirtschaftlich ist.

Dazu hieß es in der Laudatio, die Hartmut Scheffler, TNS Infratest, vortrug:

„Wenn der Innovationspreis der Deutschen Marktforschung einmal nicht an ein neues Marktforschungsinstrument ginge, das wie alle anderen auch mit Stichprobe, Befragung und der Armee der Powerpoint-Ingenieure arbeitet, sondern an ein wirklich neues Tool, das alle Marktforscher gebrauchen können.“

Ich glaube 2010 ist es soweit. Gleich mehrere Einreichungen 2010 beschäftigen sich mit Knowledge Management Systemen, früher nannte man das mal „Ablage“ oder Archiv. Das Neue an diesen für den BVM preisverdächtigen Ablagesystemen ist die Kombination mit Text Mining, das mit statistischen und linguistischen Methoden versteckte Beziehungen zwischen Texten aufzeigen kann.



Dr. Peter Gentsch, TextTech, und Stefan Althoff, Lufthansa Technik

Die Entwicklung von Insight Mining war in vielerlei Hinsicht eine Herausforderung für alle Beteiligten, unter anderem durch die Entwicklung neuer, praxisgerechter Berechnungsalgorithmen.

Stellen Sie sich mal vor, Sie geben alle ihre „Fundstücke“, also Marktforschungspräsentationen, Analystenreports, Newsletter, Zeitschriften in eine große Datenbank, im Internet natürlich. Die werden erst einmal aufbereitet, mit sprachstatistischen Verfahren der Computerlinguistik. Da werden Texte „tokenisiert“, „getagged“, „lemmatisiert“, „Disambiguierungen“ wer-

den aufgelöst und signifikante „Kookkurrenzen“ werden berechnet. Können Sie noch folgen? Auf Deutsch: Texte werden nach einzelnen Wortarten zerlegt, nach Wortarten klassifiziert und in ihre Grundform überführt, danach werden Mehrdeutigkeiten aufgelöst und das gemeinsame Auftreten zweier Begriffe berechnet.

Und was macht man dann damit?

- Zeit und Geld sparen, weil die Informationssuche vier- bis sechsmal schneller funktioniert.
- Nutzer werden automatisch informiert, wenn ein neuer Artikel zu einem für sie interessanten Thema erscheint.
- Sie erfahren automatisch, über welche Themen im Zeitverlauf stärker oder seltener gesprochen wird
- oder welche Begriffe zusammengesehen werden." ■

ALLES

IM GRÜNEN BEREICH

foerster&thelen
MARKTFORSCHUNG FELDSERVICE GMBH
BOCHUM

BUNDESWEIT UND INTERNATIONAL • IHR DIENSLEISTER
FÜR FELDDARBEIT / DATENERHEBUNG • WWW.FTMAFO.DE

UNABHÄNGIG • KOMPETENT • ZUVERLÄSSIG

Im Verbund:  www.ft-group.de · BOCHUM · DÜSSELDORF · HANNOVER · BERLIN · DRESDEN · WIEN · BARCELONA · BUENOS AIRES