

Text Mining: Wissensgewinnung aus natürlichsprachigen Dokumenten

Submitted by [witte](#) [1] on Tue, 2011-01-04 09:41

- [NLP](#) [2]
- [Text Mining](#) [3]

Title	Text Mining: Wissensgewinnung aus natürlichsprachigen Dokumenten
Publication Type	Report
Year of Publication	2006
Refereed Designation	Does Not Apply
Editors	Witte, R. [4], and J. Mülle [5]
Institution	Universität Karlsruhe, Fakultät für Informatik, Institut für Programmstrukturen und Datenorganisation (IPD)
Report Number	Interner Bericht 2006-5
Other Numbers	ISSN 1432-7864
Abstract	Das noch recht junge Forschungsgebiet Text Mining umfaßt eine Verbindung von Verfahren der Sprachverarbeitung mit Datenbank- und Informationssystemtechnologien. Es entstand aus der Beobachtung, dass ca. 85% aller Datenbankinhalte nur in unstrukturierter Form vorliegen, so dass sich die Techniken des klassischen Data Mining zur Wissensgewinnung nicht anwenden lassen. Beispiele für solche Daten sind Volltextdatenbanken mit Büchern, Unternehmenswebseiten, Archive mit Zeitungsartikeln oder wissenschaftlichen Publikationen, aber auch Ströme kontinuierlich auflaufender Emails oder Meldungen von Nachrichtenagenturen (Newswires). Im Gegensatz zum Information Retrieval geht es beim Text Mining nicht darum, lediglich Dokumente anhand von Anfragen aufzufinden, sondern aus einem einzelnen oder einem Satz von Dokumenten neues Wissen zu gewinnen, etwa durch automatische Textzusammenfassungen, die Erkennung und Verfolgung benannter Objekte oder die Aufdeckung neuer Trends in Forschung und Industrie. Durch die ständig wachsende Zahl elektronisch verfügbarer Texte werden automatisch arbeitende Verfahren zur Bewältigung der Informationsflut immer dringender, was Text Mining zu einem sehr aktiven und auch kommerziell interessanten Forschungsgebiet macht. Der vorliegende Bericht enthält eine Auswahl von Themen, die von Studierenden der Universität Karlsruhe im Rahmen eines Hauptseminars am IPD im Wintersemester 2004/2005 erarbeitet wurden. Sie reichen von den Grundlagen der Computerlinguistik über einzelne Algorithmen zur Sprachverarbeitung bis hin zu konkreten Anwendungen im Text Mining. Zahlreiche Literaturreferenzen zu jedem Kapitel sollen dem Leser eine weitergehende Studie der einzelnen Themen ermöglichen.
URL	http://digbib.ubka.uni-karlsruhe.de/volltexte/1000005161 [6]

Attachment	Size
TextMining-Bericht2006.pdf [7]	2.37 MB



Except where otherwise noted, all original content on this site is copyright by its author and licensed under a [Creative Commons Attribution-Share Alike 2.5 Canada License](#).

Source URL (retrieved on 2025-12-04 05:39):

<https://www.semanticsoftware.info/biblio/text-mining-wissensgewinnung-aus-nat%C3%BCrlichsprachigen-dokumenten>

Links:

- [1] <https://www.semanticsoftware.info/users/witte>
- [2] <https://www.semanticsoftware.info/category/topic/nlp>
- [3] <https://www.semanticsoftware.info/category/topic/text-mining>
- [4] <https://www.semanticsoftware.info/biblio/author/1>
- [5] <https://www.semanticsoftware.info/biblio/author/72>
- [6] <http://digbib.ubka.uni-karlsruhe.de/volltexte/1000005161>
- [7] https://www.semanticsoftware.info/system/files/TextMining-Bericht2006_0.pdf