

DUST-2 und Thesaurus-Entwicklungen

M. Skorsky⁽¹⁾, W. Kollwe⁽¹⁾, G. K. Hartmann⁽²⁾, A. Nölle, M. L. Richards

In Verbindung mit einem MPAE Forschungsprojekt, das die Überprüfung von Ozon- und Wasserdampfdaten beinhaltet, sollen Abstrakts vom STN international - FIZ Karlsruhe mittels der TOSCANA Software mit relevanten numerischen Daten verknüpft werden.

Problem: Im Bereich der Atmosphärenforschung zu Ozon und Wasserdampf existiert eine kaum überschaubare Menge von Veröffentlichungen und Meßdaten. Selbst für den Fachmann ist es nicht mehr möglich, den Überblick über die aktuellen Forschungsergebnisse zu behalten.
Ziel: Die Erstellung eines Informationssystems, mit dem der Nutzer schnell und zielsicher Dokumente und zugehörige Messdaten findet.
Lösung: Erstellen eines prototypischen Recherchesystems mit der Software TOSCANA unter Einsatz der Begrifflichen Wissensverarbeitung.
Vorgehen: a) Vorbereitung: Bildung des Wortschatzes für die Atmosphärenforschung durch den DFD-Thesaurus Aufbereiten des Thesaurus sowie Erstellung der Begriffshierarchien und inhaltliche Themen durch Experten. Darstellung der polyhierarchisch vernetzten Inhalte zur Unterstützung der Recherche.
 b) Werkzeug: Grafische Darstellung der Begriffshierarchien zur Recherche. Nutzung des Expertenwissen bei der inhaltlichen Analyse und Suche durch die Recherchesoftware TOSCANA, Anbindung an DUST-2-Messdaten.

„Suche Ozon-Messdaten aus der atmosphärischen Forschung. Oder so ähnlich. Gibt es da überhaupt etwas wie kann ich meine Suche eingrenzen? Welche Dokumente und welche Messdaten stehen dahinter?“

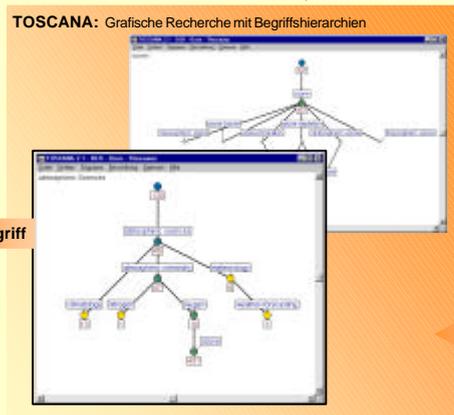
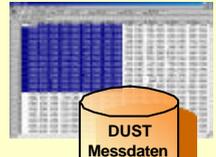
Grafisch gestützte Recherche mit TOSCANA



Antwort

Auswahl der durch die Experten vorgegebenen Themen „Ozone“ und „Atmospheric Sciences“

Recherche



In der Software TOSCANA werden die von Experten erstellten, inhaltlich zusammenhängenden Begriffshierarchien zur Recherche benutzt. Die verschlagworteten Dokumente werden den einzelnen Begriffen zugeordnet, wobei die Hierarchie automatisch berücksichtigt wird. Das Diagramm zeigt die Verteilung der Dokumente an, die sich mit „Atmospheric Sciences“ beschäftigen. Die Zahl, die unterhalb der Punkte (d.h. Begriffe) steht, gibt die Anzahl der Dokumente an, die dieses Schlagwort aufweisen. Durch die Diagramme erhält der Benutzer einen Überblick über mögliche Treffer. Funktionen des Verknüpfens von Diagrammen ermöglichen eine individuelle Eingrenzung der Ergebnismenge. Weitere Funktionen ermöglichen jederzeit den direkten Zugriff auf die Treffer.

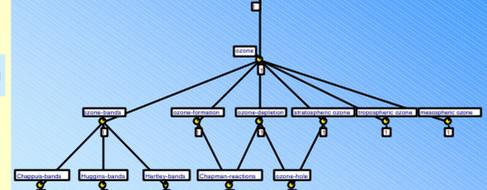
Thesaurusgestützte Begriffshierarchie

Thesaurus
 Der DFD-Thesaurus umfasst rund 7500 Einträge, mit etwa 23.000 Synonymen. Der Thesaurus deckt folgende Bereiche ab: Remote Sensing, Space Based Sensors, Satellite Missions, Space Shuttle Missions. Er ist nach den Standards der DIN 1463 aufgebaut.

Bearbeitung durch Experten:
 Analyse des Thesaurus
 Überprüfung der Hierarchie
 Zusammenstellung der Begriffe nach Themen

Pflege

Begriffshierarchie zum Thema Ozon
 Mit den Methoden der Begrifflichen Wissensverarbeitung werden polyhierarchische Begriffsbeziehungen dargestellt und in Diagrammen visualisiert. Damit ist es möglich, diese Beziehungen aus dem Thesaurus direkt zu verwenden und auch zu überprüfen.



In obigem Diagramm werden alle Begriffe mit ihren Beziehungen dargestellt, die sich im DFD-Thesaurus mit Ozon beschäftigen. Die gelben Punkte stellen die Begriffe dar und die Linien repräsentieren die Beziehungen zwischen den Begriffen. Ein Begriff ist dann Unterbegriff eines anderen, wenn er durch einen absteigenden Linienzug erreichbar ist. Er ist dann Oberbegriff wenn er durch einen aufsteigenden Linienzug erreicht werden kann. „ozone“ ist Oberbegriff von „ozone-bands“ und umgekehrt.

Nutzen des grafischen Thesaurus

- Die Dokumente müssen selbst nur fein verschlagwortet werden, Oberbegriffe werden automatisch berücksichtigt
- Strukturierung des Wissensbereiches durch den Thesaurus und die Themen
- Nutzung dieses Expertenwissens bei der Recherche mit Diagrammen
- Bessere Pflege des Thesaurus

Nutzen der Recherche mit TOSCANA

- grafische Begriffshierarchien geben Übersicht über vorhandene Treffer und erleichtern die ungenaue Suche
- Verwendung des Expertenwissens – Hierarchien sowie thematische Strukturierung des Wissensbereiches – bei der Recherche
- Verknüpfung der Diagramme und Zoomtechnik ermöglicht individuelle, themenartige Recherche
- Link auf die Dokumente ermöglicht direkte Analyse des Suchergebnisses
- zusätzlicher Link zu den Messdaten unter DUST ermöglicht die sprachliche, qualitative Recherche nach quantitativen Daten.

Die Werkzeuge: NaviCon Decision Suite

Die NaviCon Decision Suite ist ein Analysewerkzeug, mit dem der Benutzer selbständig seine qualitativen Daten mit den Methoden der Begrifflichen Wissensverarbeitung analysieren kann.

Sie besteht aus folgenden Programmen:
 - CERNATO ist das neu entwickelte Programm zur begrifflichen Strukturierung der Daten. Diese Strukturen werden in hierarchische Diagramme umgesetzt.
 - TOSCANA ist das Programm zur Betrachtung der Diagramme und zur grafischen Analyse.
 - ANACONDA ist der Diagrammeditor

Die NaviCon Decision Suite 1.0 analysiert qualitative Daten wie Produkteigenschaften, Leistungsmerkmale oder die Beschreibung von Marktteilnehmern. Die neue Lösung liefert einen vollständigen Überblick über Abhängigkeiten, Ausnahmen und Gemeinsamkeiten von Daten - also Ergebnisse, die sich durch das Verrechnen von Daten nicht erzielen und sich nicht in Zahlen darstellen lassen. Dadurch ergeben sich neben dem Einsatz als Recherchetool vielfältige weitere Einsatzmöglichkeiten:
 Im Customer Relationship Management (CRM) oder in der Marktforschung zum Erstellen von Marktübersichten und Produktvergleichen sowie zur Auswahl von Produkten anhand vorgegebener Kriterien.

Die Methode: Begriffliche Wissensverarbeitung

Der Ansatz beginnt beim Menschen: Wir denken in Begriffen und Bildern, die mit einem komplexen Muster aus Assoziationen und Beurteilungskriterien verknüpft sind. Durch die Formalisierung von Begriffen wird dieser Ansatz für praktisch umsetzbare Lösungen genutzt.

Ausgehend von Objekten und Kriterien sowie der Relation zwischen ihnen werden in der Begrifflichen Wissensverarbeitung formale Begriffe und ihre Hierarchie in sogenannten Begriffsverbänden modelliert und in Diagrammen visualisiert.

Durch die nachvollziehbare Darstellung können Beziehungen, Ausnahmen und Redundanzen in der Begriffsstruktur eines Thesaurus erkannt werden. Es entsteht eine begriffliche Strukturierung des Wissensgebietes.

Gerade bei der Recherche unterstützt diese grafische Darstellung des Wissensgebietes die oft ungenaue Suche.