science softCon

navicon

DUST-2 and Thesaurus Research

M. Skorsky[1], W. Kollewe [1], G. K. Hartmann[2], A. Nölle, M. L. Richards

In context with a MPAE research project dealing with the validation of ozone and water vapor data of the Earth's atmosphere, abstracts from STN international - FIZ Karlsruhe should be linked via the TOSCANA software with the relevant numerical data.

Problem: In the domain of atmosphere research concerning ozone and water vapor, there is an immense number of publications and measurement data. Even for the expert, it is impossible to keep all current results of research in view.

Aim: Construction of an information system which enables the user to find documents and corresponding measurement data fast and sure of aim.

Solution: Construction of a prototypical retrieval system using methods of conceptual knowledge processing and the software TOSCANA.

Realisation: a) Preparation: Creation of a vocabulary for the atmosphere research by the DFD-Thesaurus. Updating the thesaurus and creation of conceptual hierarchies and topics by experts. Support of the retrieval by representation of the polyhierarchically connected contents.
b) Tool: Graphical representation of conceptual hierarchies for retrieval. Usage of experts' knowledge within the content-oriented analysis and search with the retrieval software TOSCANA. Connection with measurement data of DUST-2.

Graphic Based Retrieval with TOSCANA

"Searching for information about atmospheric ozone – texts and data. How can we improve the selectivity of the search and combine texts with data?"

Response

Retrieval

DUST Measurement Data

Selection of the topics "Ozone" and "Atmospheric Sciences" provided by experts

Thesaurus Based Conceptual Hierarchy

Thesaurus
The DFD-Thesaurus contains about 7,500 entries, with round 23,000 synonyms. The thesaurus covers the following domains: Remote Sensing, Space Based Sensors, Satellite Missions, Spaceshuttle Missions. It is formed following the standards of DIN 1463.

Revision by Experts:
Analysis of the thesaurus
Verification of the hierarchy
Combination of concepts by topics

Maintenance

TOSCANA: Graphical Retrieval Using Conceptual Hierarchies

Documents

Access

Usage

Conceptual Hierarchy for the Topic Ozone,
Using the methods of conceptual knowledge processing, the polyhierarchical relations of concepts are represented and visualised in diagrams. Therewith, it is possible to use and check these relations of the thesaurus directly.

The above diagram represents all concepts with their relations which deal with ozone in the DFD-Thesaurus. Yellow circles represent concepts, and lines represent relations between concepts. A concept is a subconcept of another concept if it is reachable by a descending path of lines. It is superconcept of another concept if it can be reached via an ascending path of lines. "ozone" is a superconcept of "ozone-bands", and "Chappius-bands" is a subconcept of "ozone-bands".

Within the software TOSCANA, conceptual hierarchies are used for retrieval. The hierarchies are created by experts and connected with respect to the contents. The catchword-assigned documents are mapped to the different concepts, in the course of which the hierarchy is considered automatically. The diagram shows the distribution of documents which deal with "atmospheric sciences". The number beneath circles (i.e. concepts) states the number of documents to which the catchword is assigned. Via diagrams, the user gains an overview over possible matchings. Functions for connection of diagrams yield individual restrictions of the set of results. The drill-down function makes direct access to the results possible.

Use of Retrieval with TOSCANA

• Graphical conceptual hierarchies provide overview over existing matches and lighten inaccurate search
• Usage of experts' knowledge -- hierarchies as well as thematic structuring of the knowledge domain -- within retrieval
• Connection of diagrams and zooming technique enable individual topic-centered retrieval
• Links to documents enable direct analysis of the results
• Additional links to the measurement data from DUST enable language-based qualitative retrieval of quantitative data

Use of the Graphical Thesaurus

• The assignment of catchwords to documents can be restricted to the least layer. Superconcepts are considered automatically.
• Structuring of the domain of knowledge by the thesaurus and the topics.
• Usage of experts' knowledge within the retrieval using the diagrams.
• Better maintenance of the thesaurus.

The Tools: NaviCon Decision Suite

The NaviCon Decision Suite is an analysis tool by which the user can autonomously analyse his qualitative data using methods of conceptual knowledge processing.

It consists of the following programs:
- CERNATO is the newly developed program for the maintenance and conceptual structuring of the data. These structures are processed yielding hierarchical diagrams.
- TOSCANA is the program for viewing the diagrams and for graphical analysis.
- ANACONDA is the editor for the diagrams.

The NaviCon Decision Suite 1.0 analyses qualitative data like properties of products, performance attributes, or the description of members of a market. The new solution yields a complete overview over dependencies, exceptions, and common properties of data - hence, results which cannot be achieved by numerical calculations and cannot be represented by numbers.
Therefore, besides the usage as a retrieval tool, various possibilities of applications show up:
In Customer Relationship Management (CRM), in market research, and in competition analysis for the creation of market overviews and product comparisons as well as for the selection of products based on given criteria.

The Method: Conceptual Knowledge Processing

The approach starts with human thinking:
We think in concepts and images which are related to a complex pattern of associations and criteria for valuation and judgement. By the formalisation of concepts, this approach is used for practically usable solutions:

Starting with concepts and criteria and the relation between them, conceptual knowledge processing models formal concepts and their hierarchy in so-called concept lattices and visualises them in diagrams.

Because of the understandable representation, relations, exceptions, and redundancies in the conceptual structure of a thesaurus can be recognised. A conceptual structure of the domain of knowledge evolves.

Especially within retrieval, this graphical representation of the domain of knowledge supports the often inaccurate search.

Contact

Dr. Martin Skorsky [1]
NaviCon GmbH
Heinrichstr. 9
D-60327 Frankfurt/M
Internet: www.navicon.de

Dr. Wolfgang Kollewe [1]
NaviCon GmbH
Heinrichstr. 9
D-60327 Frankfurt/M
www.navicon.de

Prof. Dr. Gerd K. Hartmann, (PI) [2]
Max-Planck-Institut für Aeronomie
Max-Planck-Straße 2
D-37191 Katlenburg-Lindau
www.linmpi.mpg.de