



Sproglige problemstillinger i informationsøgning

Marianne Lykke Nielsen

Mln@db.dk

Danmarks Biblioteksskole
Institut for Informationsstudier

Viden og Dokumenthåndtering med sprogteknologi
2. December 2004

Agenda

- Informationssøgningsprocessen
- Sproglige problemstillinger
- Semantiske værktøjer
 - Tesauri
 - Topic maps
 - Taksonomier
- Konstruktionsmetoder

Søgeprocessen

- Begrebsanalyse af søgeopgaven
- Valg af informationskilder
- Formulering af søgeforespørgsel (query)
- Søgning og interaktion
- Relevansbedømmelse og reformulering af søgeforespørgsel

Søgeopgave

Søgeopgave:

Du er medicinsk reviewer. En læge har kontaktet dig. Han ønsker at få data, der beskriver kombineret anvendelse af Citalopram og Reboxetine til behandling af resistent depression. Han vil gerne se rapporter vedr. mulig interaktion.

Søgefacetter:

- Interaktion
- Citalopram
- Reboxetine
- Resistent depression

Søgeproblemer

- Forstå emner, begreber, søgetermer
- Finde søgetermer
 - Præcise søgetermer
 - Synonyme betegnelser
 - Hierarkisk beslægtede søgetermer
 - Brug- og emnemæssigt beslægtede søgetermer
- Access

Søgeproblemer

- **Diskurser og fagsprog** – det samme emne kan forstås, bruges, behandles ud fra forskellige vinkler afhængig af sociale, kulturelle, uddannelsesmæssige, arbejdsmæssige faktorer
- **Hierarkier** – det samme emne kan beskrives på forskellige hierarkiske niveauer
- **Synonymer** – vi anvender forskellige betegnelser om det samme emne, person, den samme institution, projekt, etc.
- **Homonymer** – ord kan have forskellige betydninger

Søgeproblemer

Subject crossover and overlaps

Ex.: Allocation

Special languages

- *Clinical research: Allocation* → Enrolment
- *Management: Allocation*

General language

- Allocation

Søgeproblemer

Disciplinary focus

Ex.: Service contracts

Special languages

- *Clinical research:* Service contract → Master Service agreement
- *IT:* Service contract → Service level agreement
- *Production:* Service contract

General language

- Service contract

Søgeproblemer

Semantic shifts over time

Ex.: Social anxiety disorder

Special languages

- Social phobia → Social neurosis → Social anxiety disorder

Tesaurus

Definition:

- En tesaurus er en specialordbog, der angiver konceptuel, semantisk information om begreber og termer:
 - Preferred term (PT) – den term, som man foretrækker at bruge om et givet begreb
 - Scope note (SN) – en definition, der viser hvorledes termen bruges og forstås i organisationen

Tesaurus

- Synonyms (USE/UF) – synonymer, der også kan bruges til at beskrive begrebet
 - Hierarkisk kontekst (BT/NT) – andre termer, som er hierarkisk relateret
 - Ideassociative relationer (RT) – andre termer, der er brugsrelateret til emneordet
-
- Med udbredelsen af netværksbaserede, digitale informationssystemer har tesaurusen gennemgået en kraftig udvikling m.h.t. indhold og funktion

Thesaurus news[Home](#)[Sources for keyword descriptions](#)[Search](#)[Feedback form](#)[Help](#)[RCA](#)**Analytical procedures****Description:**

The analytical procedure refers to the way of performing the analysis. It should describe in detail the steps necessary to perform each analytical test. This may include but is not limited to: the sample, the reference standard and the reagents preparations, use of the apparatus, generation of the calibration curve, use of the formulae for the calculation, etc. (ICH)

Danish:

Analyseforskrifter

Synonyms and used for:

Analytical instructions, Control methods, General analytical instructions, Product-specific analytical instructions, Special analytical instructions

Acronyms:

AP, GAF, GAF-TVC, PAF

Subject category:

L Documentation

Related terms:

H Studies, analytical methods and techniques

[Analysis](#)

[Analytical methods](#)

L Documentation

[Claims for content](#)

[Module 3 quality](#)

Broader terms:

[Documentation](#)

[Documents](#)

[Document types](#)

[Instructions](#)

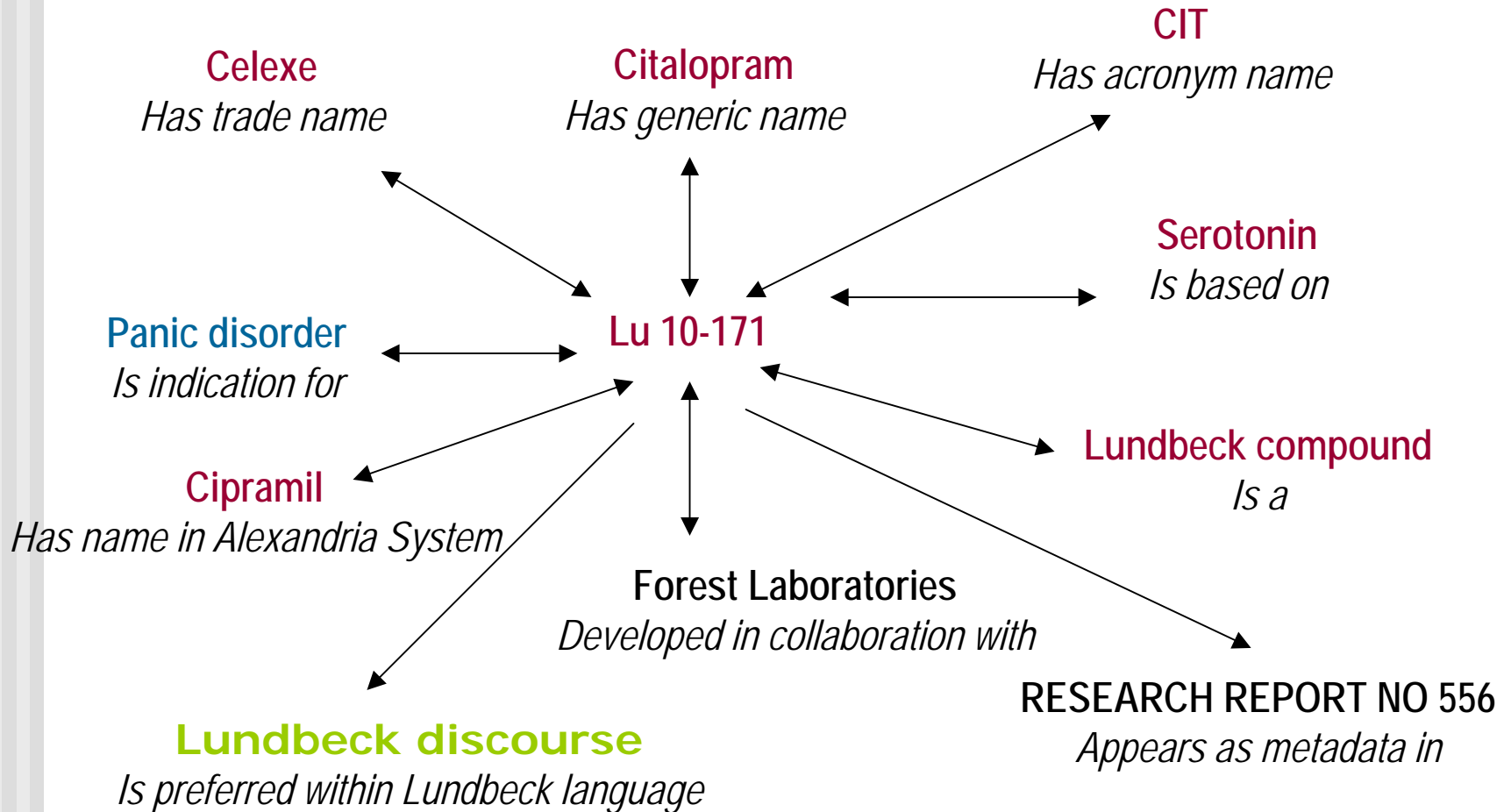
[Analytical procedures](#)

Topic map

Definition:

- Et topic map organiserer og relaterer objekter fra den virkelige verden
- En formel, eksplicit specifikation af fælles begreber - ontologistruktur
- Et topic map repræsenterer begreber og ord, men kan også repræsenterer personer, organisationer, teknologier, tekster, m.v.
- Et topic map beskriver objekter v.h.a. 'names', 'properties' og 'relationships'

Topic map



Taksonomi

Definition:

- Nyt værktøj, født i praksis verdenen
- Hierarkisk organisering af termer
- Underdeling er kontekst-afhængig og ikke nødvendigvis strengt hierarkisk
- I praksis indeholder taksonomien mere kompleks information, f.eks. synonym ring, definition og klassifikationsregler
- To grundlæggende typer:
 - Navigationstaksonomi
 - Deskriptiv taksonomi

Danfoss Konkurrentovervågning

Competitors

Financial Issues

Products

PL01 Products

Case Controller

Check Valve

Chiller Controller

Data Communication

Electromechanical Thermostats

Electronic Expansion Valves

Electronic Pressure Regulator

Electronic Temperature Controls (low end)

Filter Driers

Low and high pressure controls

Monitoring

Oil Differential Pressure Control

Oil Separators

PC Software

Powerpack controller

Pressure Regulating Valves

Pressure Sensors

Sight Glasses

Solenoid Valves

Stop Valves

Temperature Sensors

Thermostatic Expansion Valves

Water Valves

PL06 Products

PL07 Products

Refrigerants

 Find in



i Dirt remover -> Filter Driers



Search

New search

Advanced

 Search for **"Filter Driers" AND URL:pdf IN All sources**

73 Results, showing items 1 - 73

 Group by: Sort by:

Add to KnowledgeBase | Alert Me


06_FILTERS.pdf

6/15/

69 DEHYDRATORS AND FILTERS Filter shell with replaceable solid cores or filter-block for liquid and suction lines FLOW DIRECTION PESSANO -MI MADE IN ITALY Media: R22 -R134a R404A -R407C -R507 Type PS 32 bar (470 PSIG) TS -40 /+ 80 °C V L Lot N. L I S T E D 8 1 D 4 70 Castel supplies two types of

 Categories: [R134A](#) | [R22](#) | [Filter Driers](#) | [R404A](#)
http://www.castel.it/PDF/06_FILTERS.pdf - 351 KB

[Add to KnowledgeBase](#) | [Alert Me](#) | [Details](#)

24solids.PDF

Frank Garrison

7/8/

1 Many industrial drying processes call for moisture analysis as an indication of product quality or process control. An example can be found in the food industry where product specifications often include a tolerance for moisture content. The moisture content of most breads, chips, cookies, and

 Categories: [Monitoring](#) | [Sight Glasses](#) | [Filter Driers](#)
<http://www.geindustrial.com/sensing/pdf/24solids.pdf> - 28 KB

[Add to KnowledgeBase](#) | [Alert Me](#) | [Details](#)

Capa Manual de Aplicacao de Compressores.p65

Ro

2/13/

This manual may not be reproduced in whole or part without prior permission from EMBRACO. Compressor Application Manual 1 Page I. Introduction 03 II. Refrigeration circuit 04 III. Hermetic

 Categories: [R134A](#) | [Embraco S.A.](#) | [Temperature Sensors](#) | [R12](#) | [Thermostatic Expansion Valves](#) | [Filter Driers](#)
<http://www.embraco.com/ingles/produtos/00007.pdf> - 1 MB

[Add to KnowledgeBase](#) | [Alert Me](#) | [Details](#)

93402.p65

Ro

2/13/

TECHNICAL INFORMATION USE OF R 134a IN NEW LBP REFRIGERATION SYSTEMS CODE 93402 -DEC/96 -REVISION No. 02 COMPRESSOR EM 55 NP EM 55 NP EM 60 HNP Displacement cm³ 4.99 4.99 5.54 REFRIGERANT R 12 R 134a R 134a A -Evaporating Pressure (-25 C) bar 1.237 1.068 1.068 Condensing

Taxonomi

Drug products

Lundbeck compounds

Lu 10-171

Synonyms: CIT, CT, Celexe, Citalopram, Prisdal, Seralgan, Seropram,

Pharmacology

Pharmacodynamics

Mechanism of action

Nervous system agents

Central nervous agents

Antidepressants

Second generation antidepressants

Selective serotonin reuptake inhibitors

LU 10-171

Definition: Lu 10-171 is a selective serotonin uptake inhibitor. It is used in the treatment of depressive disorders

Konstruktion

Tekniske delprocesser:

- Indsamling af begreber og termer
- Formation af begreber og termer
- Organisering af begreber og termer

Ordassociationsmetoden

- Psykologisk metoden, men anvendt indenfor informationsvidenskab siden 60erne
- Simpel metoden: Stimuliord præsenteres for respondenter, der nedskriver svarord, som de de umiddelbart associerer til stimuliordet
- Anvendes til at finde relationer mellem ord, betydning af ord og sprogbrug (terminologi)
- Semi-automatisk, omkostningsbesparende metode
- Bruger- og domæne orienteret perspektiv

Serious adverse drug reactions

RT: "Bivirkning" P/BB

"Bivirkninger" P/AO

Adverse effects N/DBT

Adverse effects N/BC

Adverse event N/NO

Allergic reactions (P/MBS)

Cardiac arrest B/CT

Cardiovascular side effects B/EKM

Case report C/IBJ

Death C/HP

Disquality B/JK

Drug interaction C/KEJ

EPRS (P/MBS)

Fatal events C/KA

Fatality N/CLL

FSPV N/BC

Genetic polymorphism N/HS

Hospitalisation C/HP

Life threatening C/VP

Malreaction P/KEL

Mechanism based inhibition N/MH

Metabolite N/MH

Mutagenic effects B/EKM

Negative reaction P/KEL

OT incapacitation C/VP

Periodic safety update report C/IEKE

Ordassociationsmetoden

Ordassociationstest hos H. Lundbeck

- Udvikling af virksomhedsspecifik tesaurus
- 25 forskere fra R&D var respondenter.
Repræsenterede Basic Research, Clinical Research, Non-clinical Research og Production
- 100 stimuli ord
- 2885 svarord

Real-life test design

- Søgetest
- Slut-brugere udførte 3 kontrollerede søgeopgaver med hjælp fra test tesauri
- Sammenligning af to test tesauri
 - Associativ Thesaurus (bruger belæg), bestående af 2637 termer
 - Litterær Thesaurus (litterært belæg), bestående af a 2854 termer
- Blind test setting

Resultater

- Testpersoner var erfarne informations søgere med stor søge- og sproglig-viden
- Testpersoner valgte omhyggeligt og bevidst søgetermer afhængig af søgeopgave og diskurs
- Test personer udtrykte stor tilfredshed med test tesauri
- Ingen af tesauri inspirerede til mere interaktiv eller explorativ søgeadfærd
- Resultat er ikke overraskende, da test tesauri viste sig at være meget ens. 54% af Litterær Thesaurus blev dækket af Associativ Thesaurus

Resultater

Semantisk, konceptuel information	Ordassociations metoden	Litteratur-baseret metode
Videnskabelig diskurs		X
Domæne-specifik diskurs	X	(X)
Diversitet	X	
Aktuel information	X	

Resultater

Fordele:

- Domæne-orientering
- Hurtig, effektiv metode
- Up-to-date

Problemstillinger:

- Kræver kvalitativ analyse
- Kræver domæne viden
- Kræver stimuliord som udgangspunkt
- Kræver respondenter med viden om domænet og sprogbrug
- Kognitiv byrde

Opsamling

- Informationssøgning kræver søgeviden og sproglig viden
- Semantisk værktøjer kan hjælpe søgeren
 - Tesauri
 - Topic maps
 - Taksonomi
- Ordassociationer kan anvendes til at udvikle bruger- og domæneorienteret værktøj
- Understøtter:
 - Forståelse af begreber og termer
 - Fremfindning af begreber og termer
 - Access

Litteratur

- Nielsen, M L (2002). *Konstruktion af metadata systemer*. Arkiv. Tidsskrift for arkivforvaltning og arkivteknik, (7). 51 - 61.
- Nielsen, M L (2002). *The word association method: a gateway to work-task based retrieval*. Åbo : Åbo Akademi University Press (Doctoral dissertation). Tilgængelig på:
<http://www2.db.dk/mln/TheWordAssociationMethod/Opslag.htm>
- Nielsen, M L (2003). *Taksonomier - et nyt begreb, et nyt værktøj?* Informationspecialisten, (3). 4 - 9.
- Nielsen, M L (2004). *Conceptual tools supporting information retrieval and sharing*. Hummelshøj, M (ed). *Knowledge and Change*. Proceedings of the 12th Nordic Conference for Information and Documentation, 2004, September 1-3, 2004.