



DANMARKS
BIBLIOTEKSSKOLEN

End-user interaction with corporate digital thesaurus

Marianne Lykke Nielsen
The Royal School of Library and Information Science
Department of Information Studies
mln@db.dk

NKOS
21. August 2003



Topic

The aim is to present findings concerning user behaviour and user interaction with a corporate, digital thesaurus and discuss implication for thesaurus design.

Two user studies were carried out in a large Danish pharmaceutical company investigating searching behaviour and information seeking behaviour.

The first study formed part of preliminary system analysis determining scope and content of the corporate thesaurus. The second study made up the later evaluation of the thesaurus.



Outline

- Design of user studies
- Findings
- Implications for thesaurus design



Design of user study I

Objective

- Gain insight into the context of the information environment:
 - Work tasks, problem situations, information needs
 - Position, educational background, work and search experience
 - Information resources
 - Searching behaviour and searching problems

Data collection methods

- Group interviews, 29 respondents
- Content and discourse analysis of 50 user requests



Information needs

Work tasks/ Discourses	A priori determinability	Repetitiveness	Novelty of topic	Type and structure
Basic research	Low degree of determinability	Low repetitiveness	Known and unknown concepts	Conscious information need 1 – 2 interrelated facets
Clinical and non- clinical testing	Low to medium degree of determinability	Some repetitiveness	Known concepts	Conscious information need 3 or more interrelated facets
Sale and CI	Medium degree of determinability	High repetitiveness	Known concepts	Conscious information need 3 or more interrelated facets

THESAURUS NEWS

[HOME](#)
[Sources for keyword descriptions](#)[SEARCH](#)[FEEDBACK FORM](#)[HELP](#)[RCA](#)

Lu 10-171

Description:

Citalopram. Activity: A selective serotonin uptake inhibitor.
Indications: Treatment of depressive disorders and prevention of relapse/recurrence, panic disorder with or without agoraphobia and obsessive compulsive disorder.
(Psychotropics, Core SmPC, red) Project numbers: 104, 106, 107, 108, 111, 112, 115, 116, 117, 119, 123, 323, 335. CAS Number: 59729-33-8

Synonyms and used for:

Citalopram HBr, Citalopram HCl, Citalopram hydrobromide, Citalopram hydrochloride, Nitalapram

Acronyms:

CIT, CT

Generic (INN) name:

Citalopram

Trademarks (selected):

Akarin, Alpertia, Bufarol, Celexa, Cipram, Cipramil, Ciprex, Paranova, Elopam, Fren dol, Lupram, Prisdal, Redoxamin, Sepram, Seralgan, Seropram

Subject category:

C1 Lundbeck compounds

Related terms:

C0 Pharmacology, toxicology and chemistry, general

Racemate

C1 Lundbeck compounds

Lu 03-010

Lu 26-054

Lu 26-055

D Diseases and disorders

Agoraphobia

Depressive disorders

Obsessive compulsive disorder

Panic disorder

Prevention of recurrence

Prevention of relapse

F Psychology, physiology and biochemistry

Serotonin

M Internal and external relations

Forest Laboratories

XGMP Pharmaceutical terms

Racemate

Broader terms:

Lundbeck compounds

Lu 10-171

Pharmacology

Pharmacodynamics

Mechanism of action

Neurotransmitter agents

Serotonin agents

Selective serotonin reuptake inhibitors

Pharmacology

Pharmacodynamics

Mechanism of action

Nervous system agents

Central nervous system agents

Antidepressants

Second generation antidepressants

Selective serotonin reuptake inhibitors

Lu 10-171



Design of user study II (1/3)

Objective

- Investigation of user interaction and user satisfaction with corporate thesaurus:
 - Problem analysis
 - Query (re)formulation
 - Understanding of search concepts
 - Lead-in
- Experimental retrieval test
 - 20 researchers from R&D
 - Controlled search jobs
 - In-house retrieval system,
 - commando-based, exact match system (BRS/Search)
 - semi-controlled, manual indexing practice



Design of user study II (2/3)

Test procedure – before search sessions:

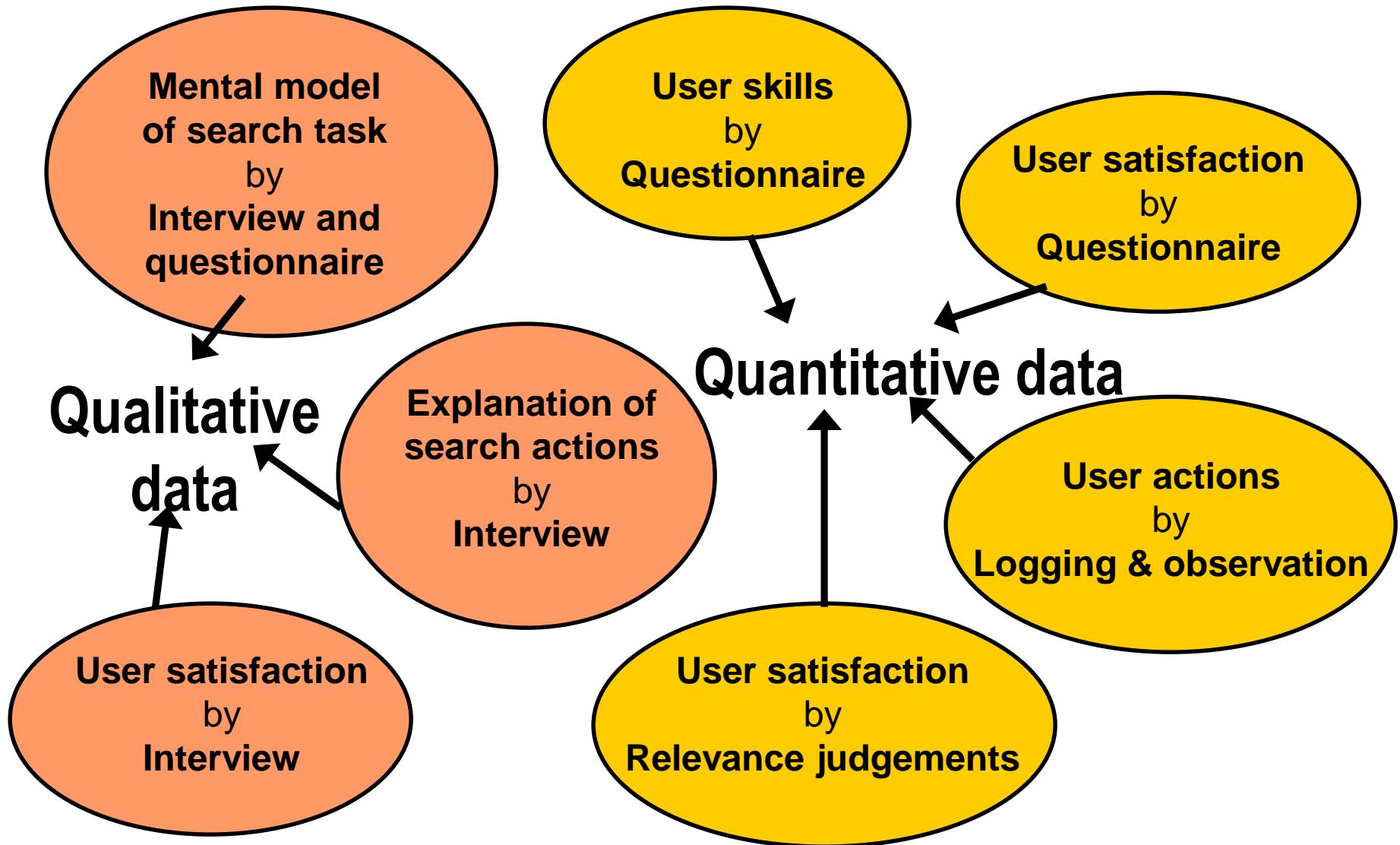
1. Investigation of searcher skills by e-mail questionnaire

Test procedure - during two-hour test session:

1. Introductory orientation session
2. Pre-search investigation of searcher's mental model of search problem and expectations by questionnaire and interview
3. Search session and investigation of user actions by observation and log-file
4. Post-search investigation of user motivation and satisfaction by interview and questionnaire

Step 2 – 4 were replicated for each search problems.

Data types and collection methods (3/3)





Thesaurus use

	(n = 58)	
Role in searching	Amount of sessions	%
Problem analysis	-	-
Query formulation	50	86
Understanding of search concepts	13	22
Lead-in	39	67



Search moves (1/2)

Search formulation moves	Amount	%
Exhaust	96	41
Specify	50	21
Parallel	47	20
Reduce	35	15
Exclude	5	2
Pinpoint	2	1
Total	235	100



Search moves (2/2)

Search term moves	Amount	%
Vary	109	73
Fix	26	18
Respell	8	5
Repeat	6	4
Respace	-	-
Total	149	100



Sources for search terms

Source	Amount	%
Search problem	190	46,6
Thesaurus	178	43,6
User interaction	38	9,3
Records	2	0,5
Total	408	100



Type of search terms

Relationships	Amount	%
Synonym (SYN)	81	45,5
Preferred term (PT)	52	29,2
Narrower term (NT)	35	19,7
Related term (RT)	7	3,9
Broader term (BT)	3	1,7
Total	178	100



Work task dependent behaviour (1/4)

	Basic research scientists (n = 32)		Clinical research scientists (n = 26)	
	Total	Average	Total	Average
Search runs	208	6,5	125	4,8
Thesaurus look-ups	115	3,6	101	3,9



Work task dependent behaviour (2/4)

	Basic research scientists (n = 32)		Clinical research scientists (n = 26)	
	Total	Average	Total	Average
Search moves				
Parallel	36	1,1	11	0,4
Vary	62	1,9	47	1,8



Work task dependent behaviour (3/4)

	Basic research scientists (n = 32)		Clinical research scientists (n = 26)	
	Total	%	Total	%
Sources for search terms				
Search problem	100	39	90	60
Thesaurus	133	51	45	30
User interaction	23	9	15	10
Record	2	1	-	-
Total	258	100	150	100



Work task dependent behaviour (4/4)

Type of search terms	Basic research scientists (n = 32)		Clinical research scientists (n = 26)	
	Total	%	Total	%
Synonym (SYN)	66	50	15	33
Preferred term (PT)	34	26	18	40
Narrower term (NT)	26	19	9	20
Related term (RT)	4	3	3	7
Broader term (BT)	3	2	-	-
Total	133	100	45	100



Summing up implications for design

- Strong domain-orientation
- Indication of discourses
- Redundancy
- Explanation
- User-controlled interaction
- Ease of use



Future research

- Importance and awareness of discourses
 - Can we "avoid" controlled indexing in in-house databases and rely on strong thesaurus informing about discourses?
- Ease of use
 - Will structured query modification improve retrieval?
- Automatic categorization
 - Will domain-specific thesaurus improve semiautomatic categorization?



Literature

Nielsen, M L (2002). The word association method. A gateway to work-task based retrieval. Åbo: Åbo Academy University Press. 342 p.

Available at:

<http://www2.db.dk/mln/TheWordAssociationMethod/Opslag.htm>