

A model of information searching behaviour to facilitate end-user support in KOS-enhanced systems

Dorothee Blocks
Hypermedia Research Unit
School of Computing
University of Glamorgan, UK

NKOS workshop - September 2004 - Bath, UK

Overview

- Scope of the model
- Background and development
- The documentation
- The model
 - Illustration
 - Using the model for different purposes
- Possibilities for further development
- Conclusions

There are notes on some slides!

Scope of the model

- Aims and intentions

- Raise an awareness of potential issues in information searching in KOS-enhanced systems
- Bridge the gap between more abstract macro level models (e.g. Ingwersen's Stratified model) and specific micro level models (e.g. Bates' tactics and Fidel's moves).
- Contribute to the reduction of problems in KOS-enhanced information searching by informing design of user support, training, interfaces and tools.

Scope of the model

- Systems and KOS

- Particularly applies to information searching systems using KOS for both searching and indexing.
- Also informs systems that use KOS for either searching or indexing.
- Main focus on thesauri – more potential due to different types of term relationships.
- Tested also with simple KOS (e.g. alphabetical lists).

Collection of empirical data

- User studies of KOS enhanced searching

2 preliminary studies

- Web: ERIC and Ovid
- Development of research questions & methodology
- Masters students and Care Science researchers

2 in-depth studies

- Stand-alone: FACET
 - HRU, University of Glamorgan
 - Investigation use of faceted thesauri for multimedia IR
- Museum and library professionals

Main findings

- End-users are capable of interacting with KOS in order to search.
- Role of KOS in information searching can be extended.
- Need for more support of the information searching process.

Collection of empirical data

- Methodology

- Sessions: training, scenarios and user-defined tasks.
- Qualitative and quantitative methods.
 - Observation, application logging, audio recording of think-alouds, screen capture, interviews, questionnaires, content analysis.
- Triangulation produced rich data on each session.
- Inductive approach resulted in identification of issues on conceptual and interaction levels.

Development of the model

- Models in the literature
 - Kuhlthau (1988) & Marchioni (1995):
Stages of information searching.
- Use of empirical data
 - FACET 1 interactions regrouped according to searching stages.
 - Basic model reviewed using data from FACET 2.
 - Risk Points documentation developed to integrate e.g. information on factors influencing decisions.

Development of the model

- Generalisation of FACET-specific interactions
- Testing
 - with existing KOS-enhanced systems & empirical data principally from the preliminary studies
 - Checking completeness and coherence
 - Refining definitions and reviewing processes

The model documentation

- Set of diagrams
 - Show how processes connect entities and decisions.
 - Accompanied by a description of the search process.
- Matrix representation
 - Developed for testing: A non-linear presentation.
- Risk Points (“RPs”)
 - Conditions and factors influencing decisions.
 - Potential causes for problematic situations.
 - Risks of behaviour with detrimental effects on the search.
 - Propose solutions to avoid/deal with problematic situations.

Propositions for applying the model

- Evaluation of system functionality
- Design of interfaces
- A framework for developing scenarios for user studies

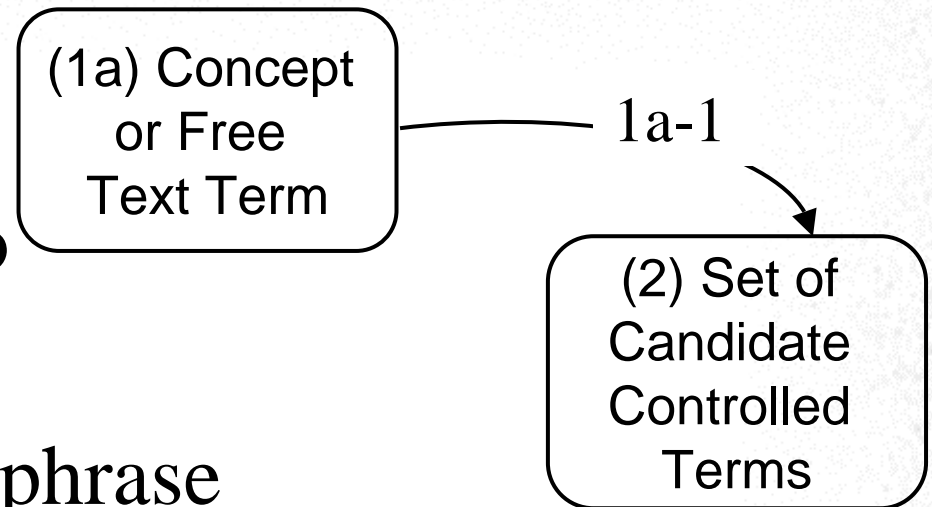
Illustrating the model

- Scenario: Starting a search
 - Mapping concepts to the KOS.
 - Selecting terms from the KOS for the query.
- Context: KOS used for searching and indexing.
- Examples from empirical studies.
- Information from Risk Points documentation.

Scenario: Starting a search

- Identifying Candidate Terms

- Concepts from the task need to be mapped to the KOS to be used in the query.



- Process **1a-1**: A word/phrase representing **(1a)** is entered into the mapping mechanism, which retrieves from the KOS entity **(2) Set of Candidate Controlled Terms**.

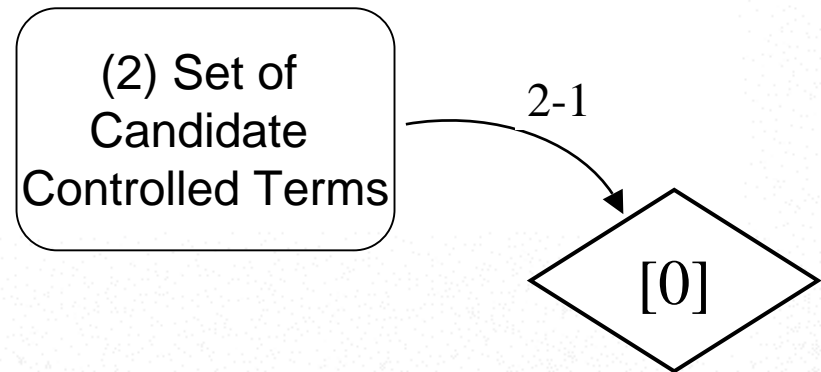
Scenario: Starting a search

- Selecting Controlled Terms

- Several possible outcomes are represented through Decision [0]: Have any Candidate terms been retrieved?
 - The number is 0 or insufficient - RP [0].1.

RP[0].1: Potential causes

- Spelling mistakes.
- Incorrect syntax.
- (Inappropriate) limitations on the mapping mechanism.



Interface design

- Is there a possibility that any of these causes and risks could occur?
- Could any of the causes and risks be avoided?
 - The number is 0 or insufficient - RP [0].1.
 - The number is too large - RP [0].2.

RP[0].1: Potential causes

- Spelling mistakes.
- Incorrect syntax.
- (Inappropriate) limitations on the mapping mechanism.

RP[0].2: Risks

- Selection difficulties.
- Limitations on the mechanism reduce number of Candidate Terms but can exclude appropriate terms from the list.

Scenario development

- Can relevant data be collected?
- How will outcomes affect the user's search, e.g.:
 - Which interaction/path will they choose?
 - Will they be able to move on to the next stage?
 - Could the evaluator assist in resolving severe difficulties?

RP[0].1: Potential causes

- Spelling mistakes.
- Incorrect syntax.
- (Inappropriate) limitations on the mapping mechanism.

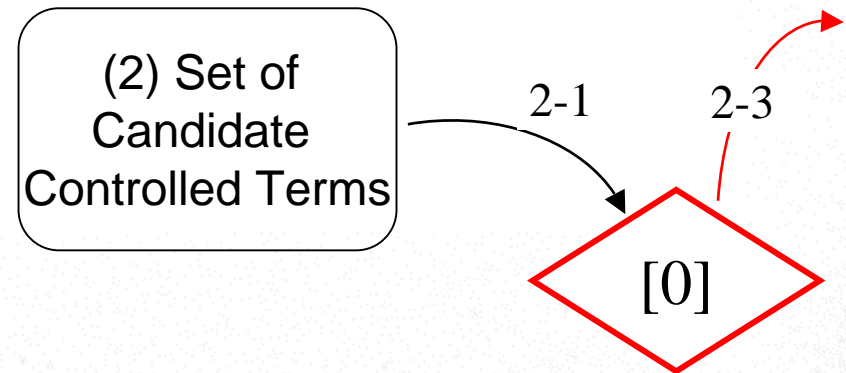
RP[0].2: Risks

- Selection difficulties.
- Limitations on the mechanism reduce number of Candidate Terms but can exclude appropriate terms from the list.

Scenario: Starting a search

- Selecting Controlled Terms

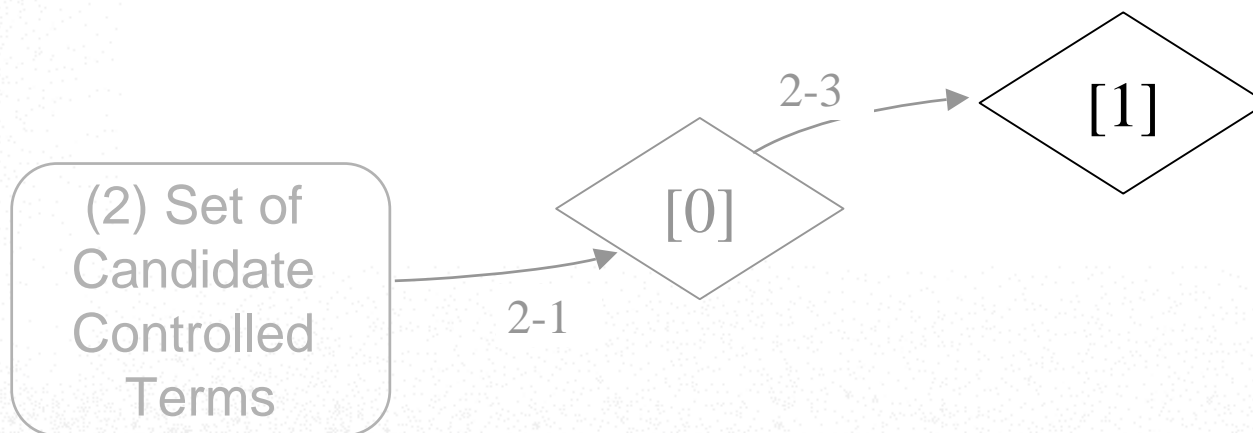
- Several possible outcomes are represented through Decision [0]: Have any Candidate terms been retrieved?
 - The number is 0 or insufficient - RP [0].1.
 - The number is too large - RP [0].2.
- The searcher starts looking for appropriate terms (**2-3**).



Scenario: Starting a search

- Selecting Controlled Terms

- Selecting of Candidate Terms ([1]) can be broken down into smaller steps not discussed here.



Scenario: Starting a search

- Selecting Controlled Terms

- Selecting of Candidate Terms ([1]) can be broken down into smaller steps not discussed here.

Interface design

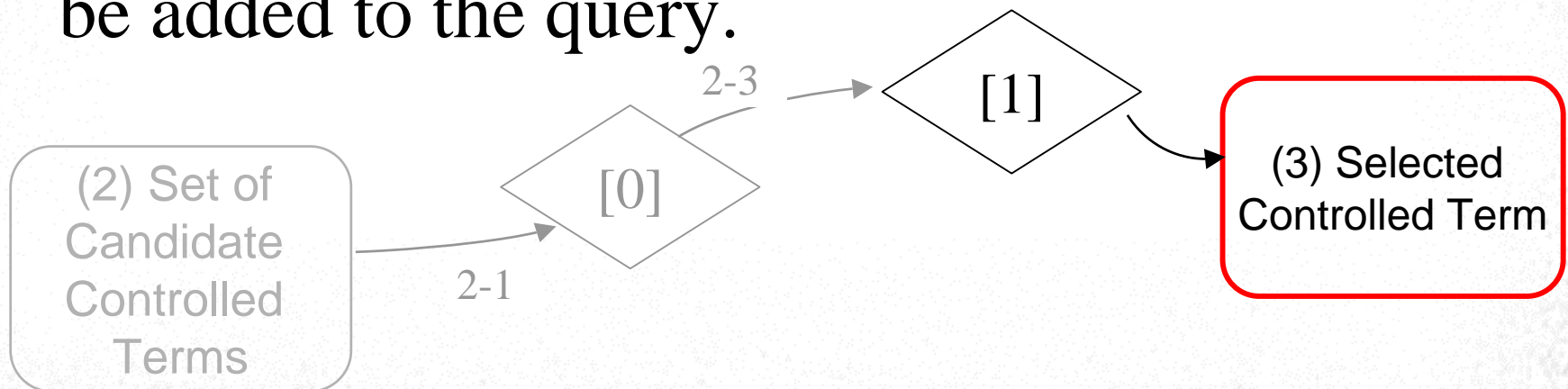
RP[1]: Some solutions/support options

- Access to terms.
- Additional resources.
- Term expansion to counterbalance non-optimal term selection.

Scenario: Starting a search

- Selecting Controlled Terms

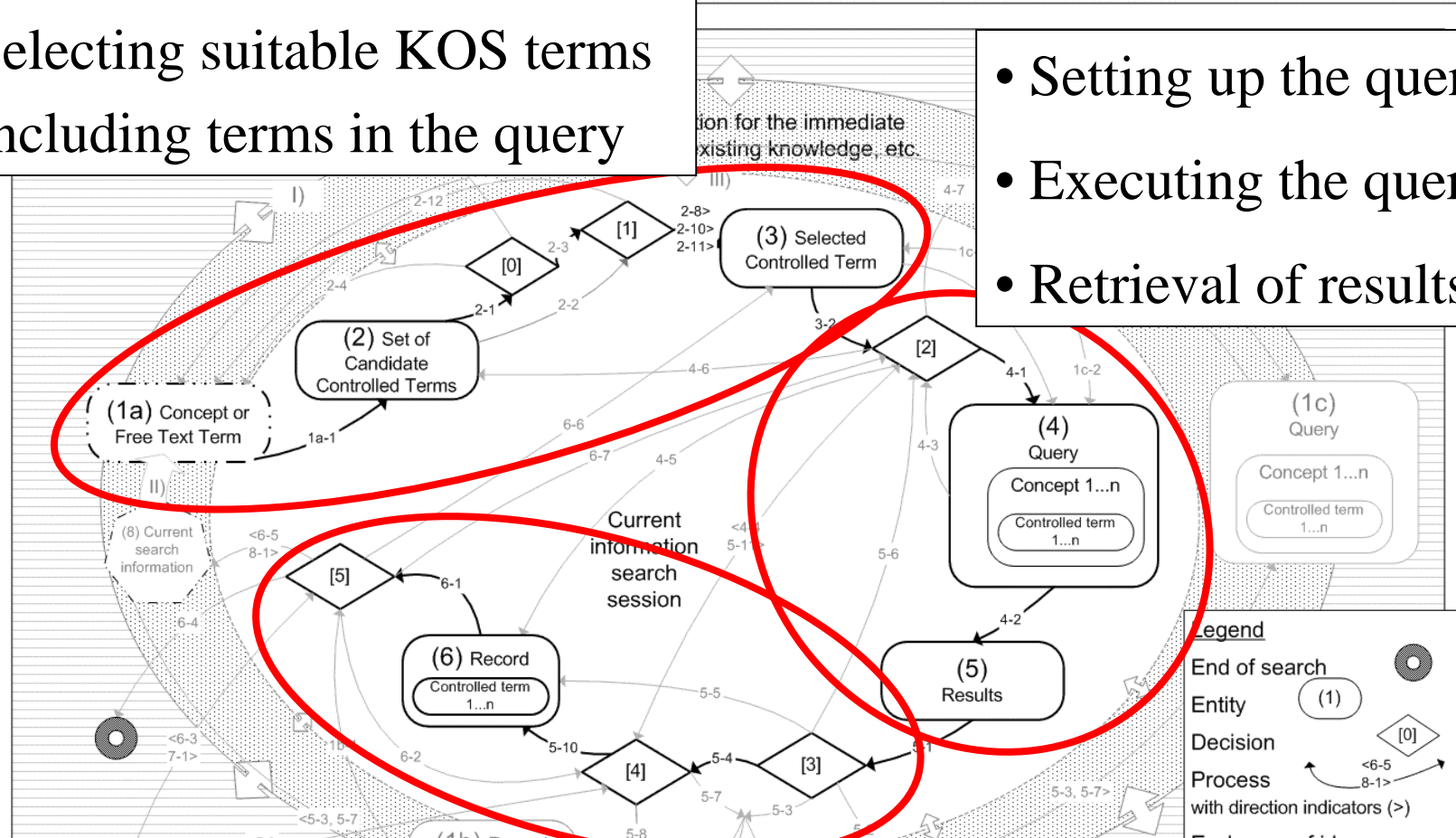
- Selecting of Candidate Terms ([1]) can be broken down into smaller steps not discussed here.
- The selected Candidate Terms are referred to as **(3) Selected Controlled Terms** and can be added to the query.



- Mapping concepts to KOS
- Selecting suitable KOS terms
- Including terms in the query

model diagram

- Setting up the query
- Executing the query
- Retrieval of results



- Evaluating the success of the query
- Inspecting individual results – using information can lead to query reformulation

Further development of the model

- Expansion to hybrid searching.
- Inclusion of more problems, risks and in particular, approaches to solutions including KOS-based tools.
- Frequency and importance of individual problems.
- (and others)

Conclusions

- This presentation
 - Overview of how the model was developed using empirical data and search stages from information searching literature.
 - Brief illustration of part of the model and how this information from the model could be applied.
- In conclusion
 - The model can inform integration of KOS for end-user searching and thus contribute to the reduction of problems in KOS-enhanced information searching through design, user support or training.

More information

- FACET project website:

<http://www.comp.glam.ac.uk/~FACET/default.asp>

- To contact me: d_blocks@web.de
- References on the next slide.

References

- Bates, M. J. (1979). Information search tactics. *Journal of the American Society for Information Science*. 30(4): 205-214.
- Blocks, D. (2004). A qualitative study of thesaurus integration for end-user searching. Ph.D. thesis. School of Computing, University of Glamorgan, Pontypridd, UK. Link at: <http://www.comp.glam.ac.uk/~FACET/publications.asp>.
- Blocks D., C. Binding, D. Cunliffe and D. Tudhope (2002). Qualitative evaluation of a thesaurus-based retrieval system. *Proceedings 6th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2002)*, Rome. Eds: M. Agosti, C. Thanos. Lecture Notes in Computer Science, Berlin: Springer. 346-361.
- Fidel, R. (1985). Moves in online searching. *Online Review*. 9(1):61-74
- Kuhlthau, C. C. (1988). Developing a model of the library search process: cognitive and affective aspects. *RQ*. 28(2): 232-242.
- Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge, Cambridge University Press.
- Ingwersen, P. (1996). Cognitive perspectives of Information Retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*. 52(1): 3-50.
- Tudhope, D., C. Binding, D. Blocks and D. Cunliffe (2002). Compound Descriptors in context: A matching function for search thesauri. *Proceedings of the Joint Conference on Digital Libraries 2002*, Portland, USA, ACM Press.
- FACET project (and more related references):
http://www.glam.ac.uk/soc/research/hypermedia/facet_proj/index.php