



Fraunhofer Institut
Sichere Informations-
Technologie

On Path-Centric Navigation and Search Techniques for Personal Knowledge Stored in Topic Maps



Outline

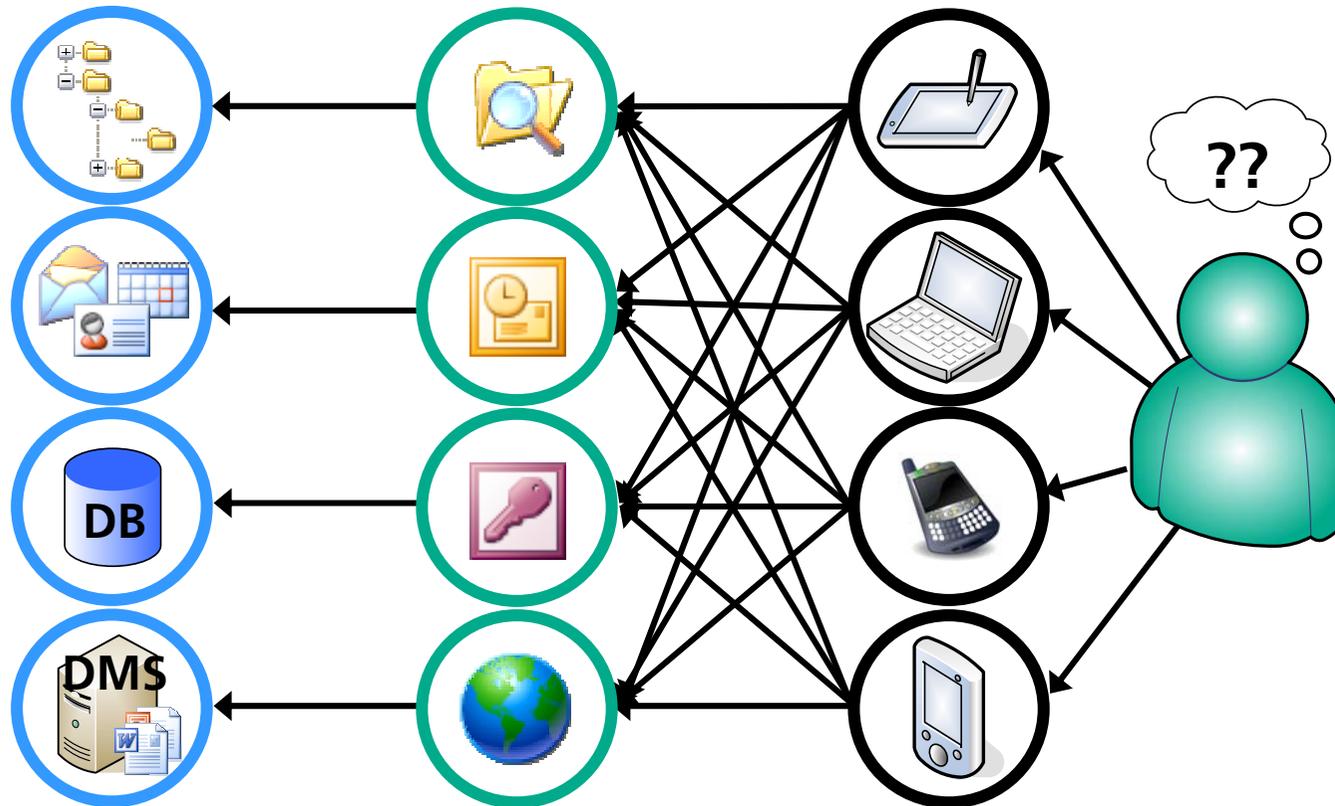
- Introduction
- Preconditions and the Model
- Navigation inside Topic Maps
- Search based on Topic Selection



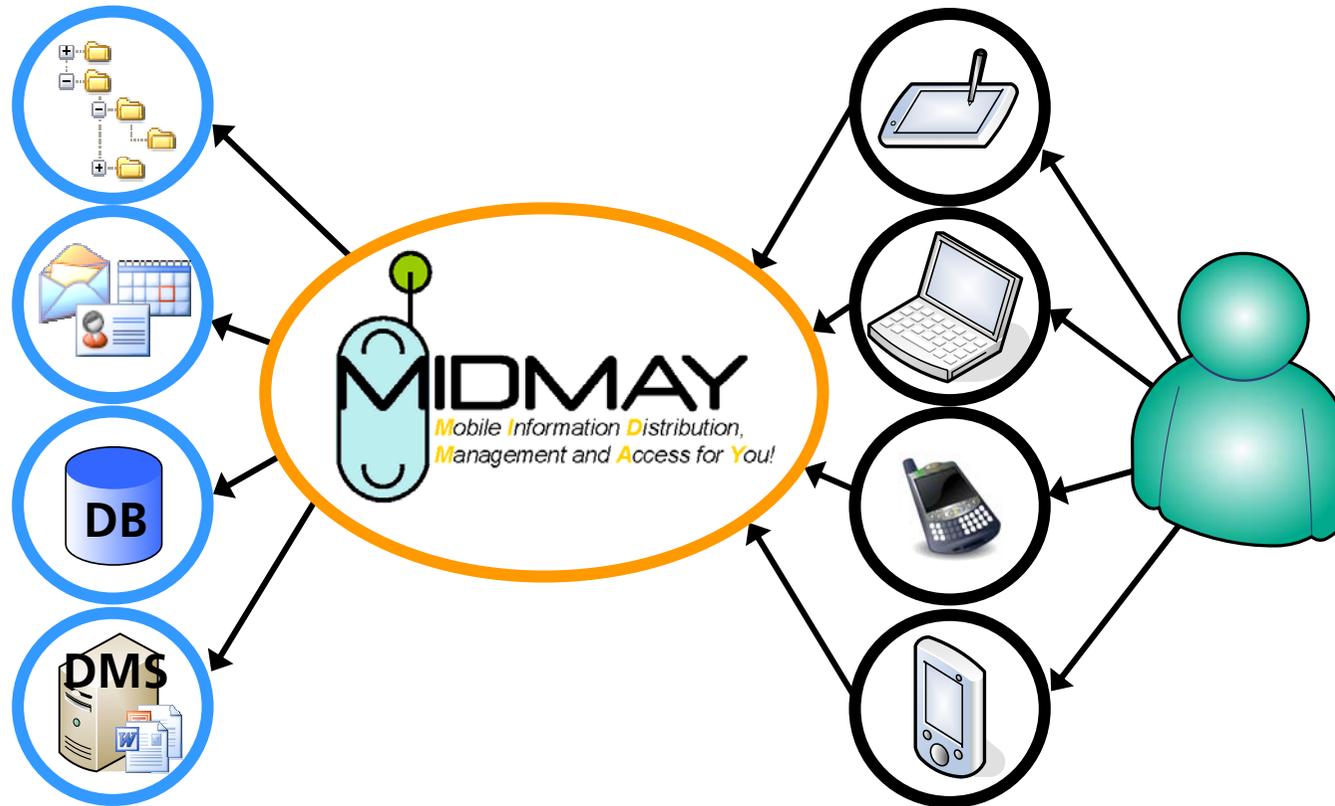
Introduction

Which problem do you address?

Daily challenges with multiple devices, various tools and different location of data



Collect and interconnect data, to ease daily work with information





Preconditions and the used Model

What is the foundation of your approach?

Topic map is designed to find the desired reference by leveraging redundancies in data sources

- Global typing schema across all extractors (PSIs, PSIDs)
- Association type reflects and unifies semantic of a property, type or hierarchical relation
- No directionality inherent in an association
- Each entry unique in the knowledge space of a user -> consistency

Graph Definition

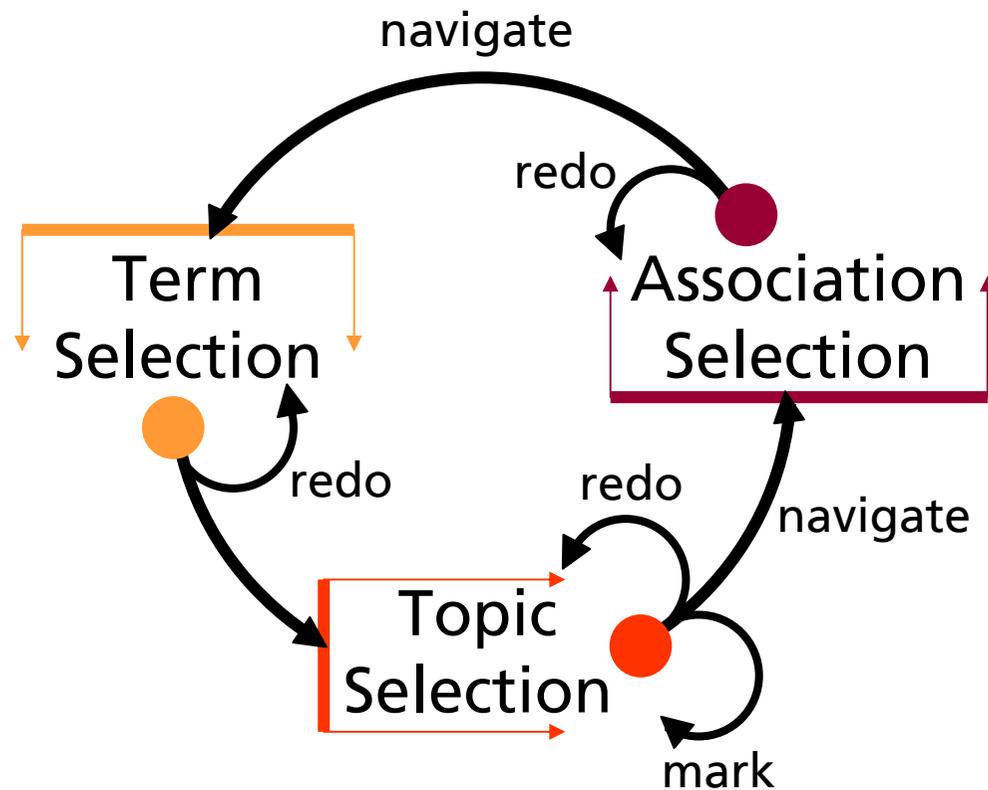
- Topic map graph G is described by the pair (V, E) .
- V is finite set of vertices mapped to topics
- E is a binary relation on V , representing the undirected associations between the topics.
- Additionally, E explicitly contains the binary relations between topics and their types $\Rightarrow G$ contains a vertex in V for every type topic.
- Each edge $(v_i, v_j) \in E$ is given a constant configurable weight w_{ij} depending on the type of association and the search mode



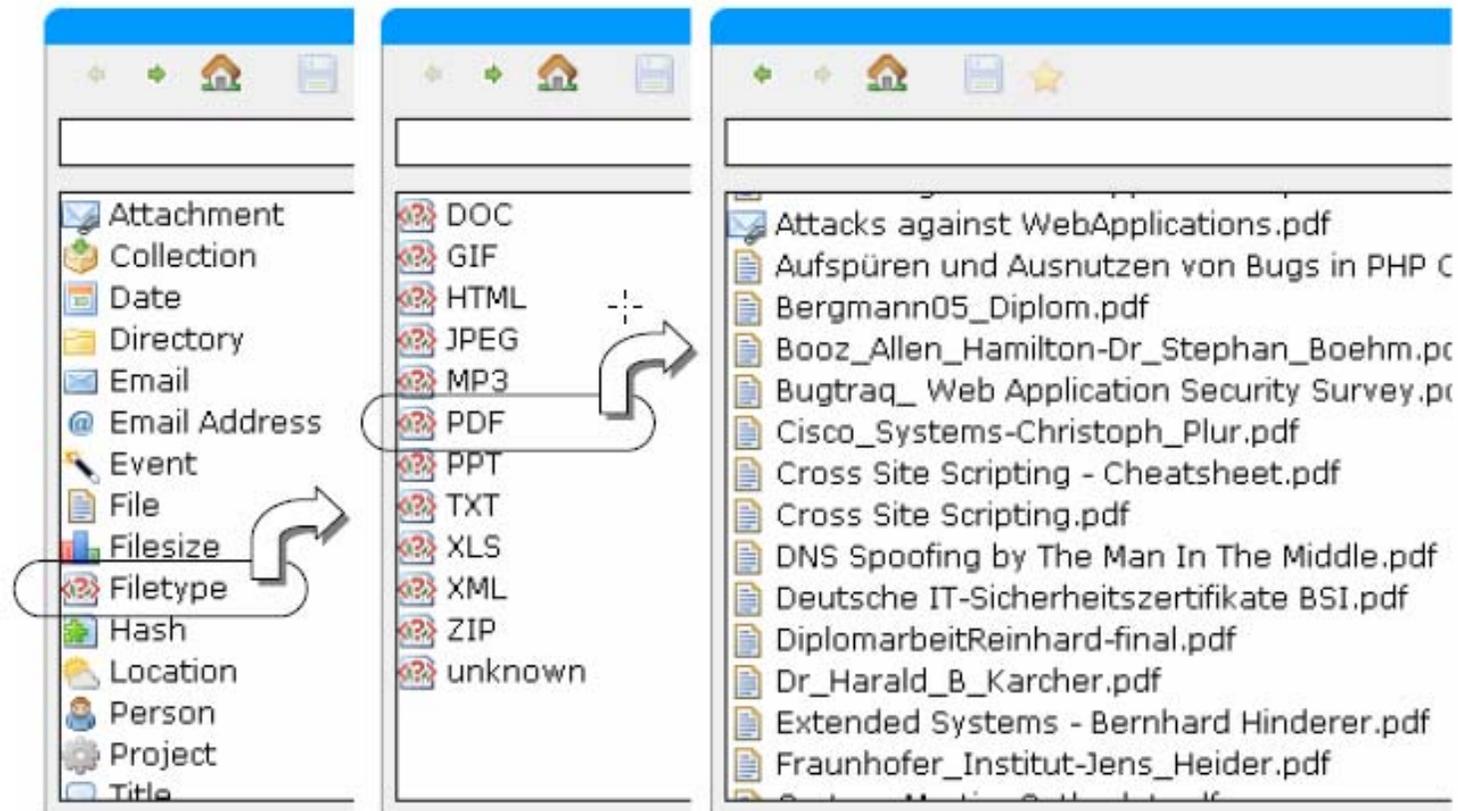
Navigation inside Topic Maps

What do you mean with path centric?

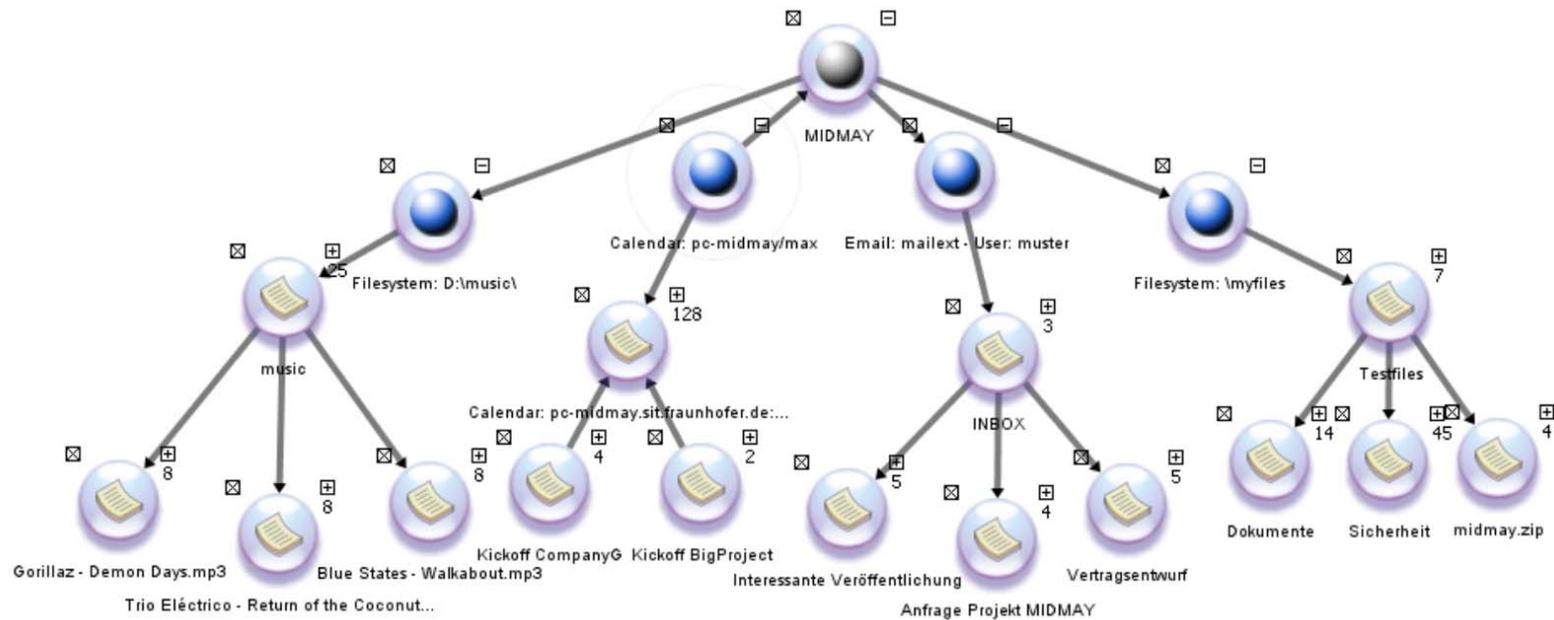
Cycling through the Graph



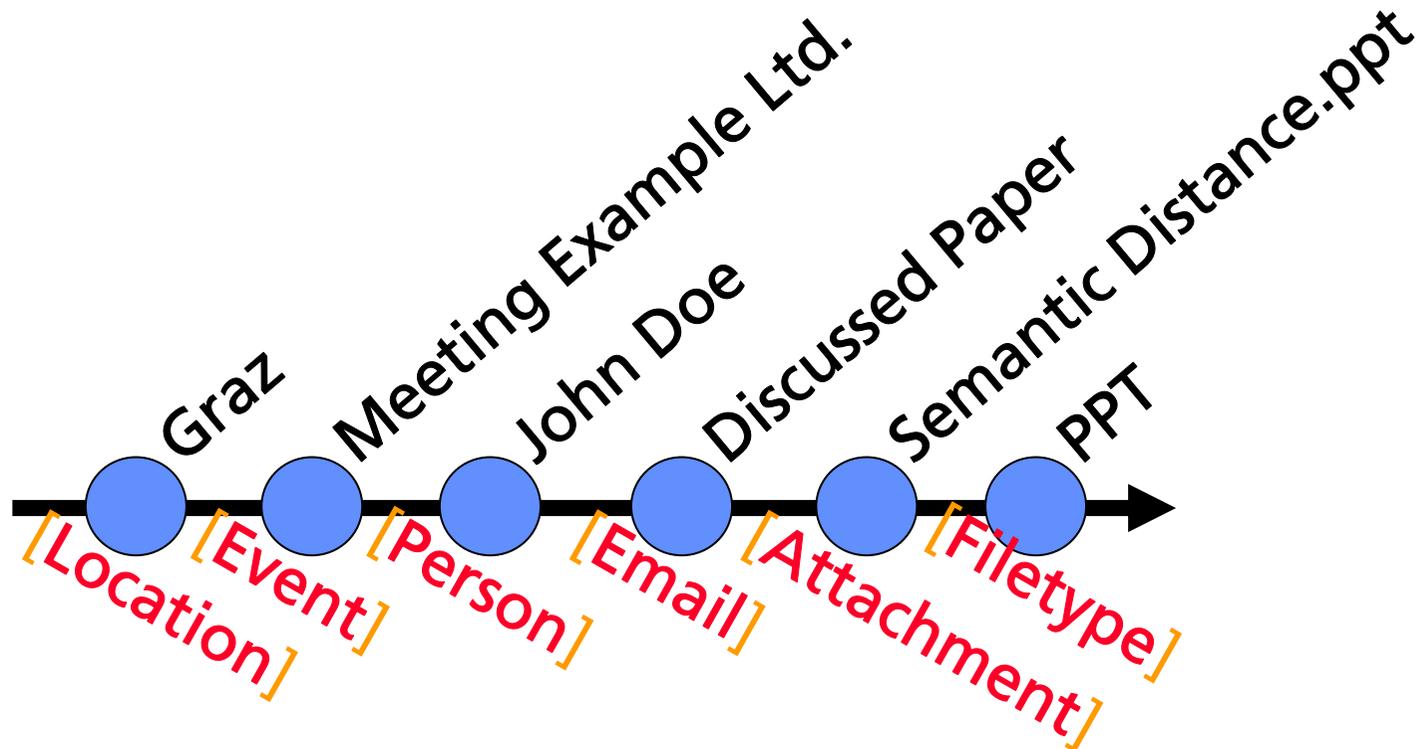
Points of Entrance & Navigation Aid Type Lists



Points of Entrance & Navigation Aid Hierarchy Root



Topic Paths





Path Navigation

- Follow paths in mind
 - I want to find a document, but can't remember some fitting keywords
 - However, I recall that the document was sent by someone I meet in a meeting in Berlin
 - Let's start the search with an item I can name: Berlin ...



Path Navigation Screenshots (1)

The image displays three sequential screenshots of a mobile application interface for path navigation. Each screenshot shows a search bar at the top with the text "Enter Term:" and a list of categories below it. The categories are: Collection, Project, Attachment, File, Person, Title, Hash, Email Address, and Location. The first screenshot shows the search bar empty. The second screenshot shows the search bar containing the text "BER". The third screenshot shows the search results for "BER", with the item "berlin[7]" highlighted in blue. The bottom of each screenshot shows a navigation bar with labels: "TermMenu", "Menu", "TermMenu", "Menu", "History", and "Menu".



Path Navigation Screenshots (2)

The image displays three screenshots of a mobile application interface, arranged horizontally. Each screenshot shows a search results screen with a search bar and a list of results.

- Left Screenshot (Topic - # 7):** The search bar contains "Enter Term:". The results list includes:
 - TU_Berlin-Torsten_Pehl.pdf
 - TU Berlin - Torsten Pehl.ppt S...
 - Presseversion_BMI-Studie_mobil...
 - Endbericht Presseversion BMI-S...
 - Berlin** (highlighted)
 - Stefan Sossna (PTP Berlin) 407...
 - Terminbestätigung Berlin
- Middle Screenshot (Association-# 1):** The search bar contains "location of". The results list is empty.
- Right Screenshot (Topic - # 4):** The search bar contains "Enter Term:". The results list includes:
 - KickOff CompanyE (04.01.2006, ...)
 - Meeting CompanyE (12.03.2006, ...)
 - Meeting Example Limited (21.09..** (highlighted)
 - Security Vortrag (11.03.2006, ...)

At the bottom of each screenshot, there are labels: "TermMenu" and "Menu" for the left and right screens, and "History" and "Menu" for the middle screen.

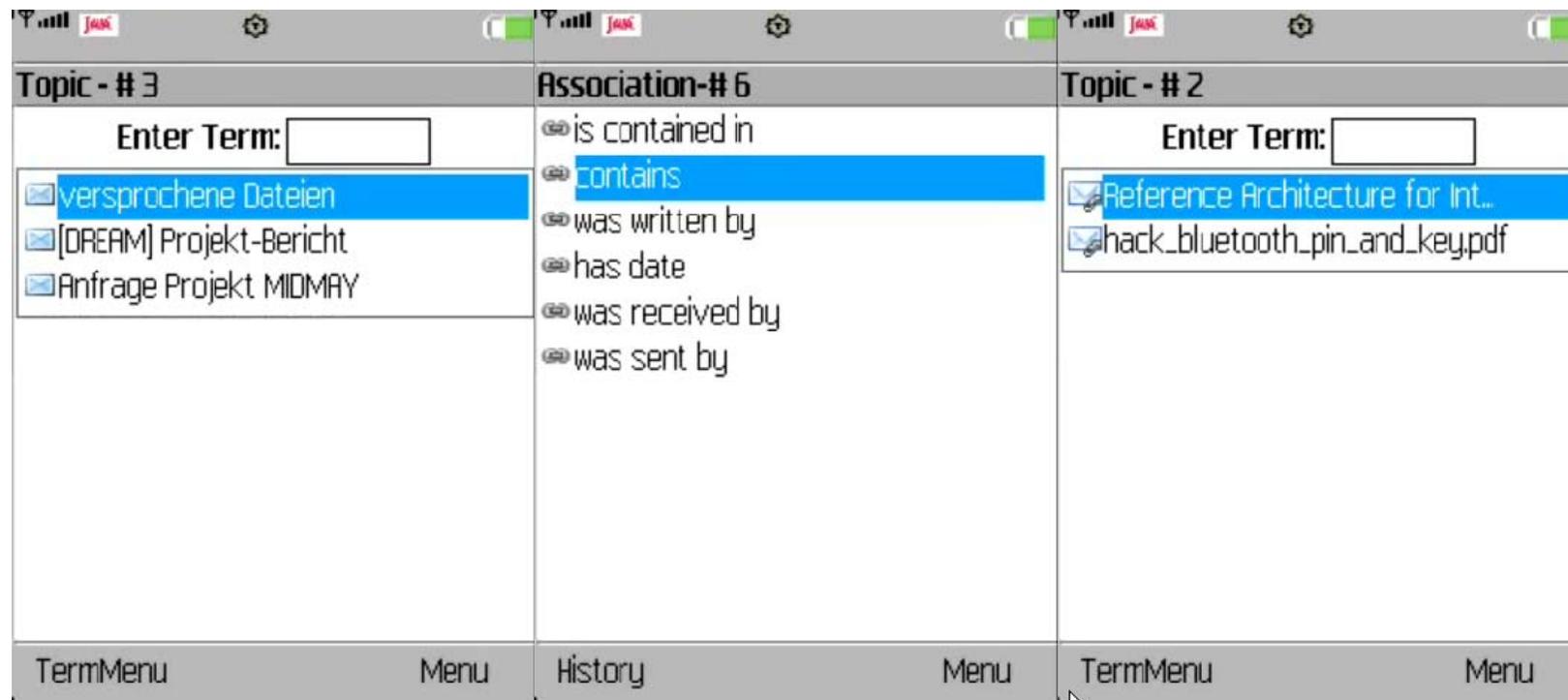


Path Navigation Screenshots (3)

The image displays three sequential screenshots of a mobile application interface for path navigation. Each screenshot shows a list of associations and a search field.

- Association-# 3:** The list includes "is contained in", "is located at", and "has attendee" (highlighted in blue). The bottom bar contains "History" and "Menu".
- Topic - # 1:** The search field "Enter Term:" contains "@erika.mustermann@sit.fraunhofer.de". The bottom bar contains "TermMenu" and "Menu".
- Association-# 4:** The list includes "is attendee of", "is child of", "belongs to", and "is sender of" (highlighted in blue). The bottom bar contains "History" and "Menu".

Path Navigation Screenshots (4)





Search based on Topic Selection

How does selecting topics help us searching?

Selecting topics to specify Search Query

- Select two topics t_a , t_b and choose mode (0-3)
- weighted Breadth-first Search from t_a and t_b , until the path with the lowest value is found
- next path by removing the edge that connects both waves -> sufficient if all topics in possible paths are presented at least once

mode	$w(h)$	$w(p)$	$w(t)$	result
0	1.0	1.0	1.0	default case; shortest paths is calculated
1	0.7	2.0	2.0	hierarchical information in result path is desired
2	2.0	0.7	2.0	path should contain interconnection of properties
3	2.0	2.0	0.7	the interests in the relation of types is expressed



Example: Show path between the topics *Graz* and *PPT*

Association path

Connections between:

Previous path Next path

Graz location of Meeting Example Ltd. (22.09.2005, 11:55-12:55) has attendee John Doe is
author of Discussed Paper contains Knowledge Synergy .ppt has file type PPT

Close this box

Path Maths Definitions introducing Bit Vectors

- The vector B_{ab}^p indicates the presence or absence of topics $\langle t_1, t_2, t_3, \dots, t_k \rangle$ in path p between topic t_a and t_b . (B_{ab} : topic presence for all shortest p)

- The OR operation between two path vectors is defined as

$$B_{ab}^p \cup B_{ab}^{p'} := \{v_j | (v_j \in B_{ab}^p) \vee (v_j \in B_{ab}^{p'})\}$$

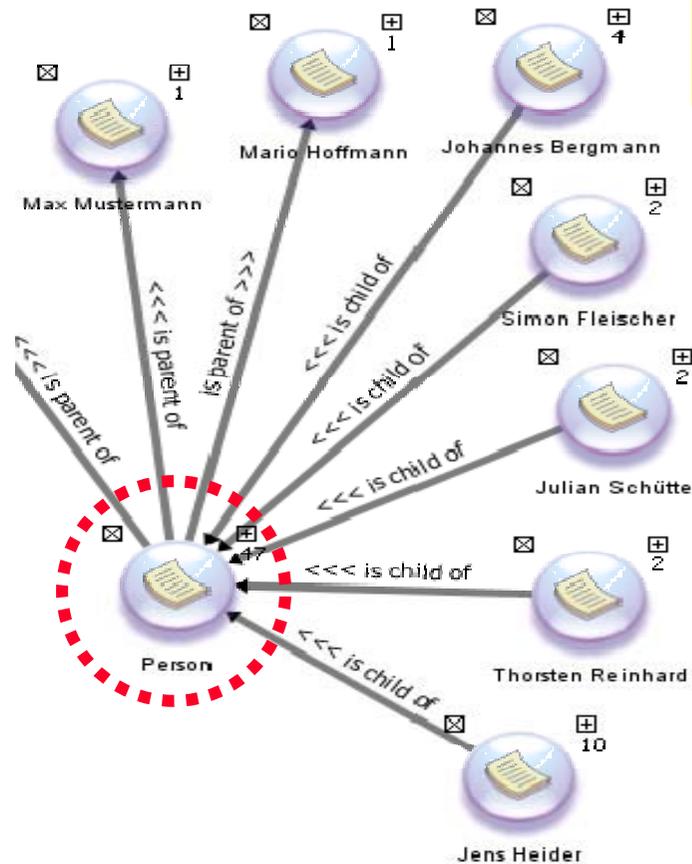
- The AND operation between two vectors is defined as

$$B_{ab}^p \cap B_{ab}^{p'} := \{v_j | (v_j \in B_{ab}^p) \wedge (v_j \in B_{ab}^{p'})\}$$

Marking multiple topics to calculate a set of relevant result topics

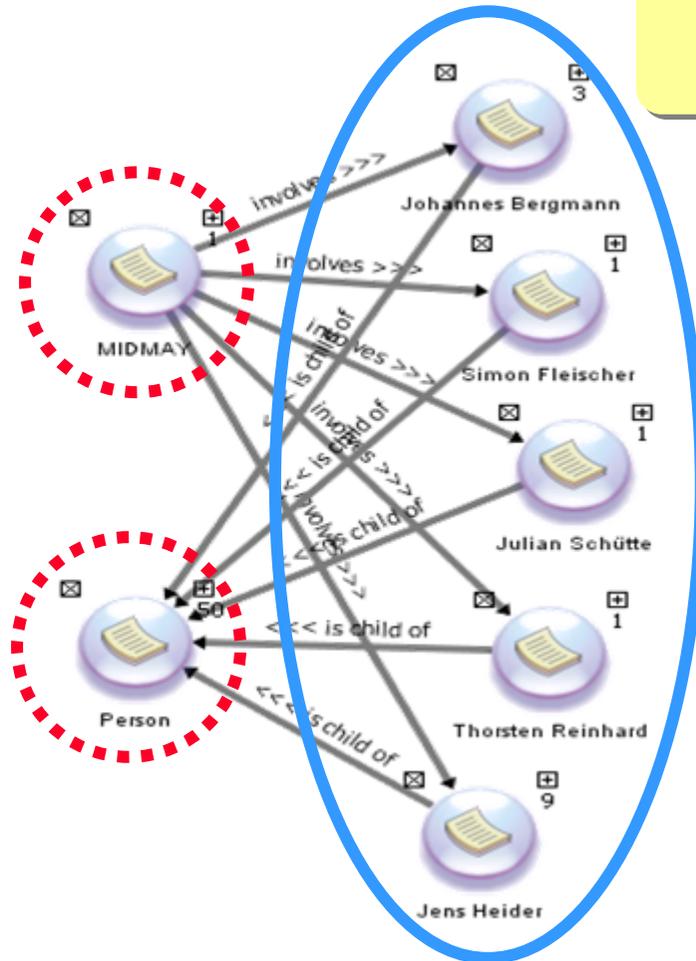
- $$B_{res}(S) := \left\{ \bigcap_{i=1, j=2}^{i=m-1, j=m} B_{s_i s_j} \mid s_i, s_j \in S, i < j \right\}$$
- Search Modes
 - Mode 0, the default mode that only uses the structure of the topic map
 - Mode 1, to focus the search on topics equally connected by hierarchy
 - Mode 2, to focus the search on equal properties of marked topics
 - Mode 3, to focus the search on equal types of marked topics

Example: Search all employees involved in project *MIDMAY* which authored a *PPT* presentation



- Think of topics related to search problem
 - I'm looking for a *person*
 - .. He's involved in project *MIDMAY*
 - The file type is *PPT*
- Navigate to the topics and mark them
- Start query

Search – Add *MIDMAY* Topic





Conclusion

What's the benefit and what are the remaining challenges?

Already existing data can be used to offer an intuitive way to search for information

- Path and set calculation provide search functionality in topic maps beyond keyword search techniques for non-technical users
- Challenges
 - Capacity of Topic Maps Engine
 - Enhanced UI for query, bringing the full flexibility of path calculation to the user
- Topic Maps can help to tackle the daily work with stored information