SIS: A system for Personal Information Retrieval and Re-Use

Summary of important points

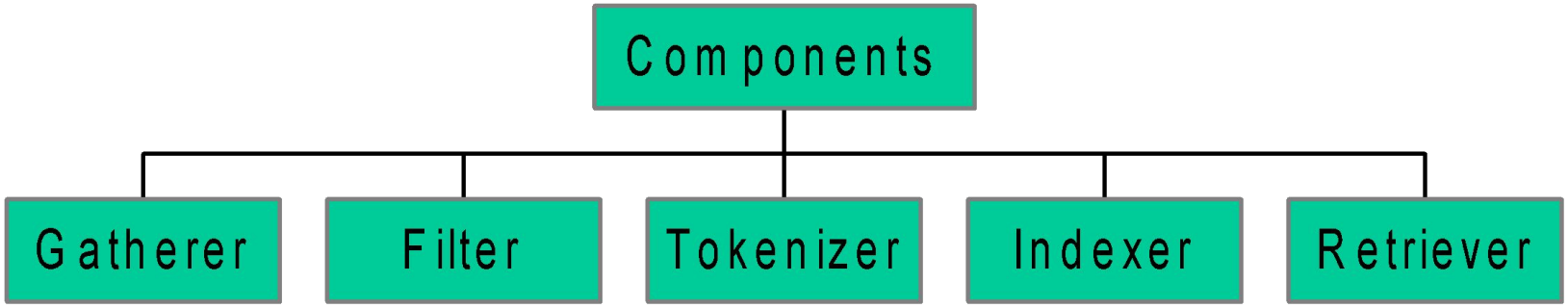
Goal of SIS

- Finding and reusing previously used information.
- Make it easy for people to find information they have already seen before.
- Increase speed of search
- Query refinement
- New ranking ideas which are more personal to user

Key aspects of SIS

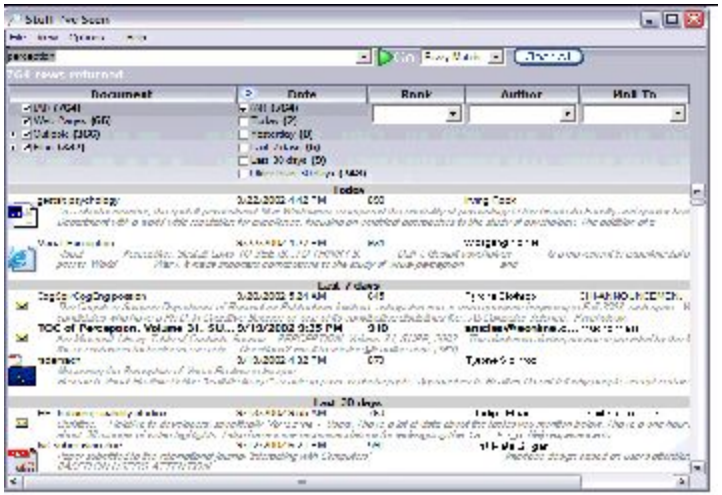
- Unified index to information from all kinds of sources on a computer
- Since information has been seen before, we have rich contextual cues obtained can be used in the searching and presenting information

System architecture



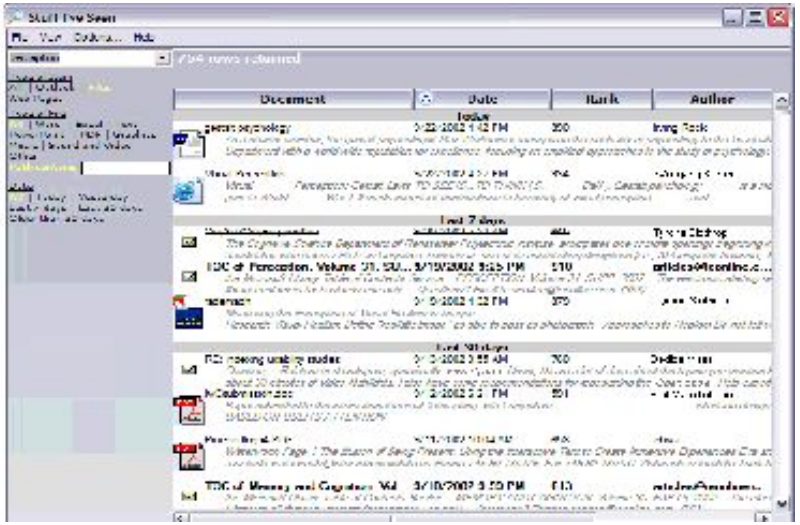
User Interface

- Top View - filters for refining attributes in each coloumn



User Interface (ctd)

- Side view - simplified filters. Eg. Outlook express



Evaluation

- Log data - Detailed information on the nature of user queries, interactions with the UI and properties of items retrieved.
- Questionnaire data - asking questions on how people organized their data before and after using SIS and about their experiences.

Observations

- 25% of queries involved people's names, indicating people are a powerful memory cue for personal content.
- Most query types were People/Places/Things, Computers/Internet & Health/Science.
- Filters used were file types and date.

Observations using log data

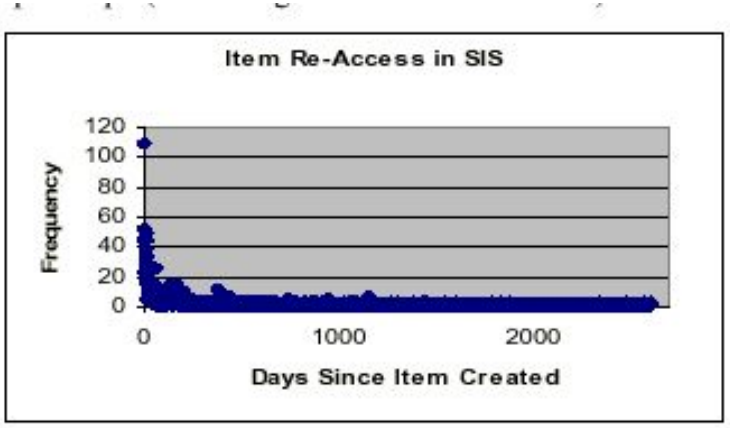


Figure 3. Frequency of access for items over time.

- Graph shows that recent items are accessed more frequently than others.

Observations (ctd)

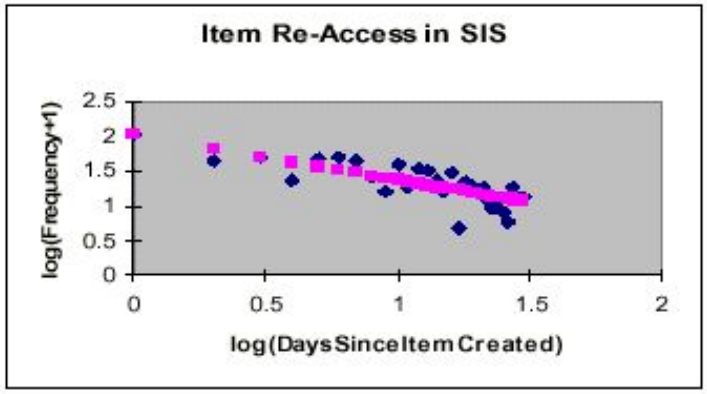


Figure 4. Frequency of access for items (log-log scale for 1 month as in [10]).

- We also see that frequency of access of items decreases since the time they are created. Email has the steepest graph since it has a shorter effective life than other documents.

User Interface observations

- Top view was preferred to side view
- Most users sorted the information by date and rank. This shows that many searches were made over personal content. Date is more useful over other attributes for sorting personal items.

Observations of questionnaire data

- The graph shows that the ease of finding information increased. Also there was a decline in non-SIS searches after people were introduced to it.

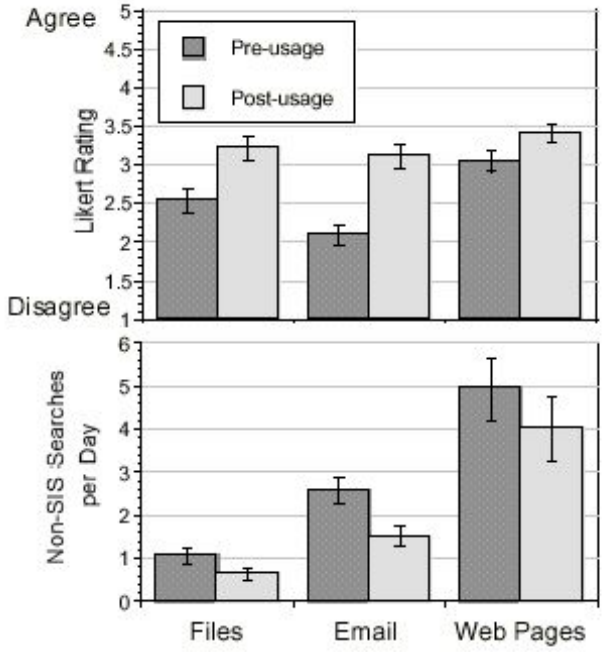


Figure 5. Ease of finding information, and frequency of use estimates, before and after SIS.

Review of paper

- No information on the implementation of their software
- More focus on the experiments and observations
- Would like more information on the unified index structure.
- Nothing novel about the user interface
- http://news.com.com/2009-1032-1020641.html MSNbot

Cognitive Expansion

- What information is relevant for expansion - sources of information are user's long term preferences, intention, situation and knowledge of specific domain
- How to use the information

Implementation

- Preference Manager - administers all explicit query preferences for each user and stores them in a suitable repository for future use
- P-Inference Engine decides what preferences to choose for expansion depending on user profile and other information sources mentioned earlier.

Architecture

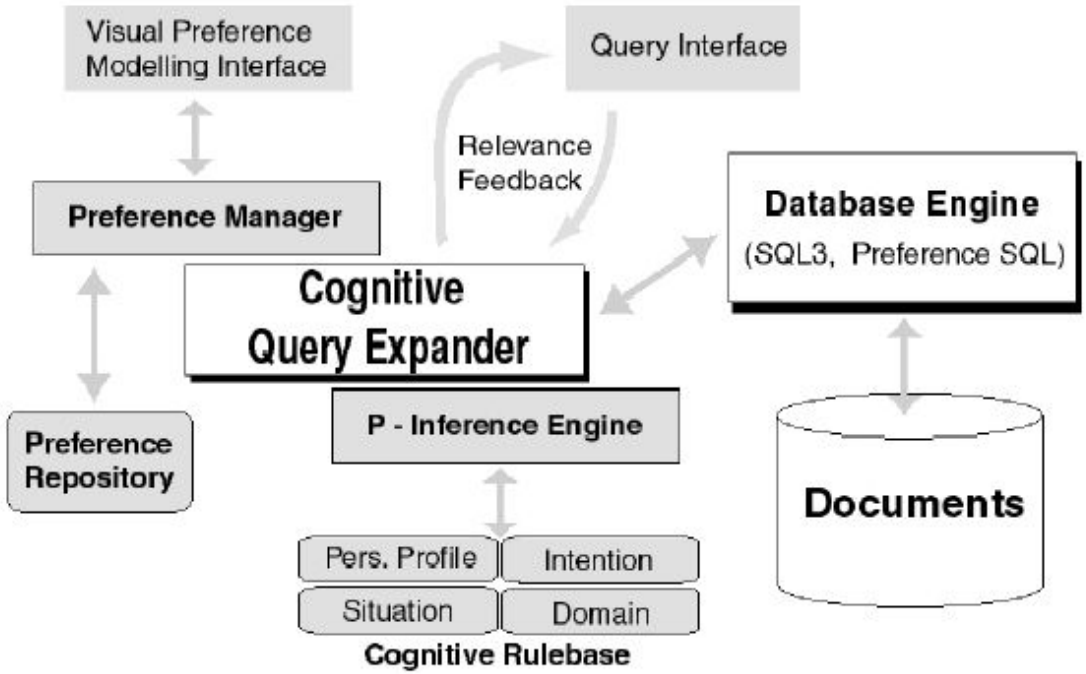


Fig. 1: Intended architecture for cognitive query expansion

Thanks!