# 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

/ Stull I've Seen				
Har kess Options				
avents.		Cin Frey, Na	🕂 💌 🕻 Jactici	2
GE reves indurinted				
Document	2 Date	Rnnk	Author	Hol To
Physics (AG4) Physics (AG4) Constants (AG4)	- Calmin (2)		Ŀ	
AH + 164.1	Lan 30 drys (5)			
	Index			
anna Hoder Pintackin X Januar Hoder Hars I	silat Loko 10 Seb 8. 10 Tenet 8. Na a montari composirine ti she eu.	or al sous persentes or al sous persentes	nchaister la a cia and	oanoth anotheria
Teste d'astronomes	Lott / de	22.	a o na Titolomo I	THANKS NOTICE
Boo Capacities Assesses Departmental conductors who have to PAU TOC of Perception, Volume And Maximum Colory Tablerd Build and Assesses School and Assesses and Assess	51. SU., 5/19/2002 9:25 PM	d coloro del mar e un dell'ordebiden Co DIC	Ab Coverante Antonio e A Inclus Vietnice - A	Rel. 200 Anh April
The second	A A A A A A A A A A A A A A A A A A A	Presidentes ( 90)	The observed of the second	<b>NR 12 Y MI</b> Externe por extensive dos
Musering by Receptor of Revents Paul Receptor of	Dir Di 2012 4 32 °M Dir 2020 4 32 °M Helle Aleximento Belgiar Int Traditto de primito de parte	en 71 SIDR, 1807	The close on close of [Am+5 = 100 -0. 80-00-01-01-010	rike z riki onaroson obriba da nyskopister zak
Annual States and Stat	Distanti en distanti en distanti en Distanti en distanti en distanti Peter Andrea en descenti Peter Statistica en distanti en descenti en distanti Distanti 20 di	na 20.000 and Paderses, 990 17	Паладорона адлагараа Гранн Siz IIVC - Полагон Солособ Со	in <mark>er i na s</mark> anakalka dar Manang kanakalka dar
Horner Hanner	Consideration of the activation of the 20 States of the States of the Provide States of the States of the States of the States of the States of States of the States of th	ner, 27, 2029, 2027 Production (1970) Induction (1970) Inductio	The distance in a second (assessor the second secon	ngen an en

## 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
- <u>http://news.com.com/2009-1032-1020641.ht</u> <u>ml</u> 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!