Organizing and Managing Personal Information

By Neha Gupta
Outline

- What is Personal Information?
- Motivation
- Related /Existing Work
- Work in progress
- References
- Conclusion
What is personal information?

- Emails
- Documents
- Addresses
- Notes
- To dos
- Photos & Videos
- Music

Any information you choose to call personal.
Motivation

- Information is scattered in different applications and devices.
- Do you sometimes spend hours navigating and searching for the files?
- Do you forget where you filed/saved your documents?
Goal

- To minimize the resources - time, energy and money spent in organizing and later reusing the information.
- To get the right information, at the right time, at the right place and in the right form.
Age of Information Overload

- Are we living in an age of information overload?
- Are we being bombarded by far too much information than we can assimilate?
- Psychologists claim that too much information can lead to stress and confusion.
Information Fragmentation

Related pieces of information are being handled by different applications and tools which leads to the problem of information fragmentation.

- Folders - for documents
- Email Client - Outlook/Apple Mail
- Address Book
- Calendar
- iTunes / Windows Media Player
- iPhoto
- Facebook / Twitter / Blogs
Filing Income Tax Returns

- Store your W-2 forms sent by your employer.
- Savings Bank Interests statements.
- Home Loans Information.
- Charity statements provided by the respective organizations.
- Securities and Stocks Returns.
Desktop/PC Searching …

- Google Desktop Search
- MAC OSX Spotlight
  - They can search within the documents, web pages visited, emails sent/received.
  - Very fast way to retrieve information.
  - No need to navigate through the hierarchical structure of folders.
Don’t take my folders away: Organizing personal information to get things done

– A study was done in which 14 participants were interviewed.
  - Participants included 4 professors, 2 support staffs, 2 graduate students and 2 librarians
  - Rest of the people were professionals in engineering.
– The interviewers were given an informal definition of project.
– Interviewers were asked the following questions:
  - Why they create folders?
  - If they are given the option of something like Google Search, can their folders be taken away? Why or why not?
Only 1 of the 14 people agreed to part away with folders given an option of Desktop Search.

Folder provides a structure to the information. Advantages of folders:
- Trust
- Control
- Visibility

What is missing:
- No support for ordering
- Tension between current reuse
- No support for reuse structure
*Picture from Reference 1*
Taking Email to Task

- Using email as the central tool for Information Management.
- Whole PIM revolves around email.
- Emails are made-
  - Task centric /Threading
  - Equality of content
  - Meta information is stored
  - Aggregated view of information -warnings,actions.
  - Users provided a positive feedback.

- Disadvantages
  - The entire PIM cannot be organized according to the emails.
  - No performance numbers were given.
Project Based-The Universal Labeler

- Labels are created according to the users action plan.
- Documents, emails can be dragged and dropped in the labels.
- The problem of Information Fragmentation can be solved.
- Disadvantages:
  - Everything in your desktop cannot be categorized as a project.
    - How do you manage audio/video and pictures which are not related to any project?
    - A random dinner with your school friends, its pictures, emails???
Organizational committee
  • Ask Hank for suggestions.
  • Contact Sally.
Conference hotel
  • Get bids from downtown hotels.
  • Check out the hotels and bed & breakfasts closer to campus.
Conference dates
  • Get dates that work for members of the org committee.
Invited participants
  • Get suggestions from committee members.
Conference web site
  • Look at other conference web sites for ideas
  • Hire a web master
  • How much money can we spend?
Conference theme

*Picture from Reference 3*
Continued...

*Picture from Reference 3
Common Underlying Representation

- **Haystack**: Any piece of information is called a URI. Uses RDF (Resource Description framework) to provide meta-data. Users can annotate, classify and relate pieces of information together.

- **Disadvantages:**
  - How to import everything into RDF?
  - Which ontology to use?
  - Performance of the system
Semantic Desktop

- Term was first coined by Stefan Decker.
- It is a desktop extension for the semantic web.
- Each file/document is identified by a URI and data can be queried through RDF.
- Social Semantic Desktop includes Semantic Desktop, P2P and social networking.
- Eg. If a group of people share some documents, then the meta-data or the annotations created by one person should be accessible by all.
Information scraps

- Scribbled Personal Information
- Post-it notes, corners of notebook, put in text files
- Information scraps form an essential part of our lives
  - Serial numbers of shipments
  - IP addresses of servers
  - Recipes
  - Directions
  - To dos
  - Reminders
Needs and Design Affordances

- Lightweight Capture - Physical/Mental Barrier should be less
- Flexible contents and representation
- Flexible use and organization
- Visibility and Reminding
- Mobility and Availability
Lotus Agenda and Chandler

- Free form information management
- Categorization of data and related actions and conditions could be made.
- “See Monica at 11:00 in Stamp Union”
- A particular item could be categorized in different ways and time line as well as priority could be set up.
- Problems:
  - Can handle only textual information
  - Cannot import data from MS word documents and internet
Jourknow

- System developed by Haystack group at MIT.
- User can create and tag structured and unstructured notes.
- Additional contextual information is automatically added to the notes like location, time, other activities on the machine.
Picture from Reference 4
Apple’s Mobile Me

- User is provided with seamless access to his files and information.
- Documents/Files, Calendars, emails, address book are kept in the cloud.
- All the information is kept in sync.
- Their tag line:
  - “Everything up to date. Anywhere you are”
- Can be used only with Apple products.
- Not free :(}
Till now...

- We see that a lot of tools have been proposed in the area of Personal Information Management.
- A consolidated approach is required to overcome hardware/software and physical/mental barriers.
Model of Personal Information

- All the different pieces of information can be called information items.
- Personal Information is defined as a collection of information items and the inter-relationships between these items.
- Meta Data, which is information about information, is also considered an information item.
Context plays a very important role in human behavior.

A particular information item may be important for a person in one scenario but may not be important in another scenario.

Study by Barreau\(^8\) shows that every user may place different importance on the same item.

So can we design a desktop unique for a person?

According to a study by Kwasnik\(^9\), the way people organized information depended heavily upon the context.

- “on the top shelf are books that are very seldom used”
- “correspondence I must deal with immediately goes to my briefcase”
Information Tags

- All the information items are associated with multiple classification tags and action tags.

**Classification Tags:**
- All information items are classified into groups.
- Eg. Current Document/Presentation can be tagged as
  - CMSC818g
  - Spring 2009
  - PIM

**Action Tags:**
- Action Tags are associated with information items and can be used to trigger certain events.
- Eg. Action can be deletion of the information item from the system.
  - Each information item has a sunset time/time of relevancy.
  - Eg. Email saying “Lets meet tomorrow at noon”, becomes obsolete after tomorrow.
Classification Tags

- Multiple classification tags are put on an item.

- Single information item falls in different categories providing more ways to identify it as compared to the folder structure.

- A hierarchical structure is put on the tags.
  - Eg. A reply of an email message inherits the tag of the email message along with its own tags.
  - A document with higher version number acquires the property of its predecessor.
Action Tags

- Can be of different types:
  - Alarms
  - Sunset Time/Time of Relevancy.
  - Triggers can be set to send email /reminder mails.
- Helps the user to clean the system.
  - Prevents information overload.
- Automatic task management
Auto-Tagging

- Is the backbone of the entire system.
- Users do not want to spend their time tagging information.
- Physical/Mental Barrier has to be reduced.

- Example Techniques:
  - Latent Semantic Analysis
  - Using TF-DF (Term Frequency Document Frequency)
  - Using temporal characteristics.
  - Converting data into semantic form.
Personal Information

- URL
  - Contains
  - Created By
- Sender
- Receiver
- Calendar Entry
  - Contains
- Documents
  - Contains
- Timestamps
  - Contains
  - Received By
Design goals

- Flexibility in organizing and later using information items.
- Low Barriers - Ease of Use
- Consistency Maintenance.
- Information Visualization.
- Low Performance overhead.
Open Questions /Problems

- **Information Visualization**
  - How can we enhance the users experience without hiding/dwarfing the main content?

- **Performance of the Personal Computing System**
  - The proposed system should not over-burden the machine and make other operations slow.

- **Auto-Classification** should be as accurate as possible and different methodologies need to be tested.

- **Keeping Information in sync** like Mobile Me and Microsoft Groove.
CONCLUSION

- This is just the starting of our thought process in the area of PIM
REFERENCES

1. Jones, W., Phuwanartnurak, A; Gill, R; Bruce. Don’t Take My Folders Away: Organizing Personal Information to Get Things Done.

2. Belloti, V., Ducheneaut, N., Howard, M., Smith, I. Taking email to task: the design and evaluation of a task management centered email tool.


9. Kwasnik, B. The role of classification structures in reflecting and building theory.
THANK YOU & QUESTIONS :)