Folder Structure & Information Retrieval

Personal Information Management in Your Everyday Life
Personal Information Management and Folder Structure

- Personal Information Management (PIM)
  - Grew out of studies on knowledge acquisition
  - Began to enter into academic research in the late 80’s early 90’s
  - Sought to study and describe individual behaviors such as cognition, technology interactions, productivity, and more

Personal Information Management and Folder Structure

- Folder Structure

  - Topic of particular interest to PIM researchers

  - Important implications for how individuals arrange information and how this impacts access / recall

  - Offers insight to private info management as well as organizational info management
Personal Information Management and Folder Structure

- What exactly is folder structure?

  - Organization of electronic information

  - Uses folders to arrange via certain organizational criteria

  - Varies by individual in terms of organizational approach/model

  - Usually discussed in terms of breadth, depth, granularity, and their impact on access/recall
Personal Information Management and Folder Structure

- Research models
  - Both qualitative and quantitative approaches

- Observation of individual behavior
  - Screen shots, over-the-shoulder observation, screen recordings, self-reporting

- Research focuses on...
  - Navigational practices
  - Time spent navigating
  - Hierarchy depth vs. breadth
  - Behavior prediction
  - Conceptual organization
  - And more…
Folder Hierarchy of Joseph Koivisto’s Personal Computer
(October 2013)
Discussion

- Your own folder structures
- How do you retrieve information you’ve saved?
- How do you arrange folders on your computer?
- Any tips? Organizational horror stories?
Important Publications

- Journal of the American Society For Information Science & Technology
- SIGCHI Bulletin
- SIGCHI Conference Proceedings
- ACM Transactions on Intelligent Systems and Technology
Studies

- Present on five studies
- Chronologically arranged
- Illustrate changing research focus over time
Barreau & Nardi (1995)
Finding and Reminding: File Organization from the Desktop

- Frequently cited as a ground-breaking study
- Important to PIM in general, folder structure studies in particular
- Studied sample groups of DOS/Windows users and MAC users
  - Back when there was a huge difference in tech and users
Barreau & Nardi (1995)
Finding and Reminding: File Organization from the Desktop

- Filing and finding – one of the most frequent computer activities
  - Divided between “location-based finding” and “logical finding”

- Different OS provide some file finding functionality
  - DOS: /whereis
  - MAC: Find utility

- Overall, both sets of users prefer location-based finding
  - Navigating through folders rather than keyword searching
Implications

File names are not most relevant type of identification

Placement of object within a (sub)folder provides contextual information about...
- File Identity
- File Use

Retrieving & Reminding are intertwined
Investigating how individuals conceptually and physically structure file folders for electronic bookmarks

- Study to explore how users create structures to organize sample set of objects

- Studied students in a Masters of Management and Professional Accounting program

- Asked to organize 60 websites in a hierarchy consisting of...
  - Superordinate folders – only contains folders
  - Primary folders – contains both objects & folders
  - Subfolders – only contains objects
Results

On average…

2.75 Superordinate folders
12.93 Primary folders
2.86 Subfolders
Investigating how individuals conceptually and physically structure file folders for electronic bookmarks

- Trends emerged from folder analysis and interviews
  - Similarity of high level folders based on disciplinary terminology
  - Subjective understanding of folder function
    - Even among identically named folders
  - Vague understanding of conceptual distinctions
  - Despite variance of hierarchy, contextual factors remain paramount
    - Relevance to current work, intended use of object
Boardman & Sasse (2004)
“Stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management

- **Study arranged to identify…**
  - Concepts used to name folders within computer folders, email application, and web browser
  - Amount of conceptual overlap between the three

- **Documented the structures of 31 participants**
  - Male-biased sample group
  - University researcher biased (29)
“Stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management

<table>
<thead>
<tr>
<th></th>
<th>Number of Folders</th>
<th>Average Folder Depth</th>
<th>Average Unfiled Items</th>
<th>% Collection unfiled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Files</td>
<td>56.6</td>
<td>3.3</td>
<td>66</td>
<td>3%</td>
</tr>
<tr>
<td>Email</td>
<td>32.3</td>
<td>1.7</td>
<td>828</td>
<td>41.6%</td>
</tr>
<tr>
<td>Bookmarks</td>
<td>16.8</td>
<td>1.3</td>
<td>44</td>
<td>38.8%</td>
</tr>
</tbody>
</table>
Boardman & Sasse (2004)
“Stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management

- Classified subject of folders and identified most frequently used organizational criteria

- File Folders
  - Project; Document Class; Role

- Email
  - Role; Project; Contact; Topic; Mailing List

- Web Browser (Bookmarks)
  - Topic; Document Class; Project; Contact
Boardman & Sasse (2004)
“Stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management

- Findings indicate a high degree of conceptual overlap between three platforms
  - Mainly between “Project” and “Role”

- Retrieval techniques of all three platforms similarly oriented
  - Browsing and navigating most frequently used techniques
  - Biased towards recently accessed materials (currently relevant to work)

- Perceived value of information promotes good filing behavior

- Despite all these similarities, storage practices remain idiosyncratic among different platforms
  - Centralized multiplatform storage
  - Parallel folder structure
Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy (2010)
The effect of folder structure on personal file navigation

- Modeled study to capture quantitative information in order to develop statistically identifiable trends

- Focus of study included...
  - Identifying depth, size, internal structure of folder hierarchies and identify relation between structure, depth, and subfolder distribution
  - Factors impacting success rate of folder navigation
  - Relation between
    - Folder depth & retrieval time
    - Folder size & retrieval time
    - Folder depth, size, and success rate
Average folder depth of active files: ~ 3
  - Most folder hierarchies do not go deeper than 4 subfolders to retrieve current information

Average folder size: 22 files
  - As depth increased, file number decreased

Average subfolders: 10.64 folders
  - Increased number of subfolders indicates conscious construction and may indicate presumption of information retrieval (and the expectation of necessary reminding)
Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy (2010)
The effect of folder structure on personal file navigation

- 94% success rate
- Average navigation: 14.76 seconds
- Average # of steps (i.e. folders accessed): 4.4
- 3.72 seconds/step
Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy (2010)
The effect of folder structure on personal file navigation

Retrieval Time = 4.956 + (2.236 X Depth) + (0.106 X Size)
Prior studies have found that increasingly complex hierarchy requires more time to navigate.

Translation: MORE MONEY

FolderPredictor attempts to automate folder navigation by using scoring system applied to recorded behavior.
Multi-Tasking Hypothesis:
1. All users break work into discrete units to which they give names – here called tasks
2. At any moment, users work on a single task at a time
3. Knowledge of current task reduces uncertainty in predicting what a user is trying to do

Weighting mechanism
- Recency
- Frequency
After extensive testing...

- Reduced click (step) costs by 49.9%
- 90% of FolderPredictor navigations place user in hierarchy within 3 clicks of desired objects
Conclusion

- Studies of folder structure investigate the varying ways in which folder hierarchy...
  - Is instrumental to information retrieval and use
  - Is guided by user organizational criteria and concepts
  - Presents opportunities for efficiency through predictive modeling

- The future of folder structure studies
  - Increased predictive programming
  - Better understanding of (un)conscious elements that influence structure development

- Raising consciousness of folder structures is important
Bibliography


