Personal Digital Document Management

Sarah Henderson
Department of Information Systems and Operations Management
University of Auckland

s.henderson@auckland.ac.nz

THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Introduction

- Personal Digital Document Management

  - the process of acquiring, storing, managing, retrieving and using digital documents

THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Files overloaded Mars Probe

Nasa scientists say hundreds of computer files that have accumulated on the Mars rover Spirit may be the cause of problems that have crippled it.

These "cruise files" will now be deleted from the second Mars rover Opportunity before it rolls on to Mars to begin its science mission.
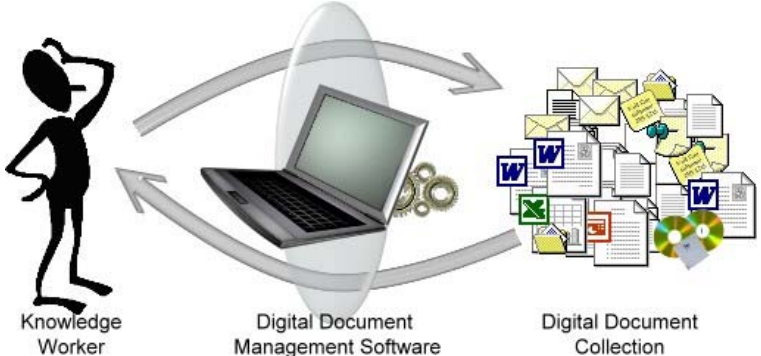
Engineers knew the probe was alive

Monday 26th January, 2004 http://news.bbc.co.uk/1/hi/sci/tech/3431617.stm

THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Personal Digital Document Management

Knowledge Worker   Digital Document Management Software   Digital Document Collection

THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Introduction

- Documents are now primarily created, stored and communicated electronically

- Anecdotal evidence suggests personal document management software (ie Windows Explorer) is not very usable

- Improvements are challenging:
  - Embedded nature of document management activities
  - Hard to imagine alternative systems
  - Individual differences

THE UNIVERSITY
OF AUCKLAND
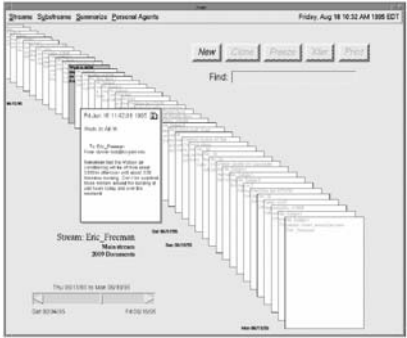NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

People can organise things by

- **Time**

- Space

- Time and Space

- File Type/Application

- User-defined attributes

- Labels/tags

Lifestreams (Freeman & Gelernter)

THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

People can organise things by

- Time

- **Space**

- Time and Space

- File Type/Application

- User-defined attributes

- Labels/tags

Data Mountain (Robertson)

THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

People can organise things by

- Time

- Space

- **Time and Space**

- File Type/Application

- User-defined attributes

- Labels/tags

TimeScape (Rekimoto)

THE UNIVERSITY
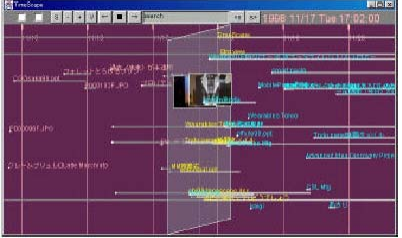OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

People can organise things by

- Time
- Space
- Time and Space
- **File Type/Application**
- User-defined attributes
- Labels/tags

Window XP Task Bar (Microsoft)

caltex.doc - Microsoft Word	
hosting interfaces.doc - Microsoft Word	
mobil.doc - Microsoft Word	
Profiling.doc - Microsoft Word	
DeclarativeSecuritySupport.doc - Microsoft Word	
C# Language Specification.doc - Microsoft Word	
CHINZ 2005.doc - Microsoft Word	

7 Microsoft Of... ▾   4 Microsoft Offi... ▾   EN

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

People can organise things by

- Time
- Space
- Time and Space
- File Type/Application
- **User-defined attributes**
- Labels/tags

Presto/Vista (Dourish, Edwards, LaMarca & Salisbury)

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

People can organise things by

- Time
- Space
- Time and Space
- File Type/Application
- User-defined attributes
- **Labels/tags**

Gmail labels   &   Flickr tags

▼ Labels
christmas list
family
grand canyon
househunting
party (3)
school (2)
vacation plans
            edit labels

You're looking at all the public photos tagged with **sunset**.

Related:
sky, **clouds**, **beach**

See also:
sun, water, sea, tree, ocean, trees, landscape, lake, silhouette, blue, reflection, dusk

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Metadata in the hierarchical folder system

Folders ✕
- Desktop
  - My Documents
    - Courses
      - 2005
        - INFOSYS 222
          - Lectures

Name ▲
- Lecture 12 Notes.doc
- Lecture 12.ppt

- Setting up a hierarchy of folders is essentially equivalent to pre-defining a set of attributes or keywords that can be applied to a document.

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

File System Snapshot

- Software has been custom-written to gather information about the file system
- Supplements the subjective accounts
- Analysis gives:
  - Number of files and folders
  - File type profile
  - Depth and breadth of hierarchy
  - File age profile
  - "Hotspot" analysis

- Folder names and structures
- File names and types
- Date created, last accessed and modified

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
Sarah Henderson, s.henderson@auckland.ac.nz

File System Statistics

	A	B	C	D	E	F
Years Experience	3	3	3	10	10	10
Files	3,790	20,327	3,793	1,545	3,021	3,614
Folders	228	2,710	854	211	415	609
Empty Folders	31	539	50	7	9	72
Files per Folder	16.6	7.5	4.4	7.3	7.3	5.9
Average Depth	2	6	6	4	6	4
File Duplication %	11.7%	76.4%	23.8%	18.3%	36.0%	28.5%
Folder Duplication %	12.7%	81.8%	42.5%	37.9%	33.5%	51.0%
Shortcuts on Desktop	35	4	9	5	36	12

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
Sarah Henderson, s.henderson@auckland.ac.nz

Desktops (Colour coded by number of files)

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
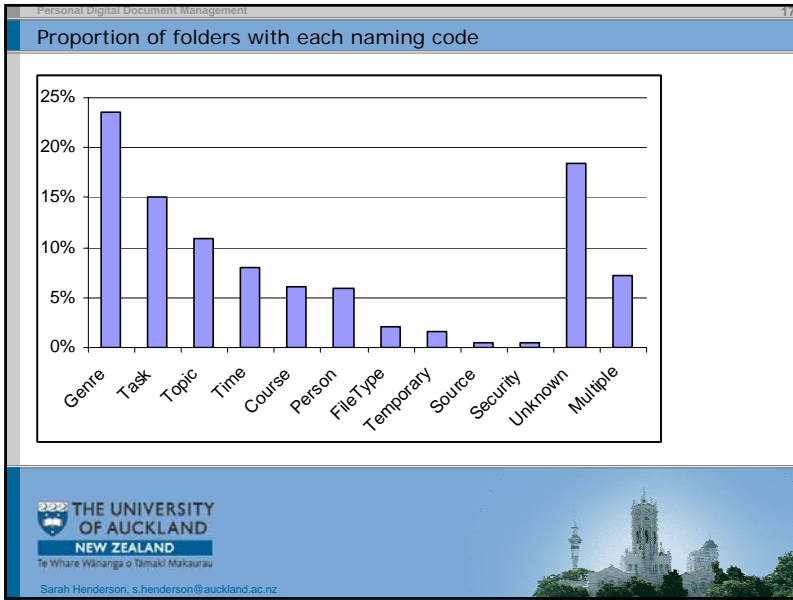Sarah Henderson, s.henderson@auckland.ac.nz

My Documents (Colour coded by number of files)

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
Sarah Henderson, s.henderson@auckland.ac.nz
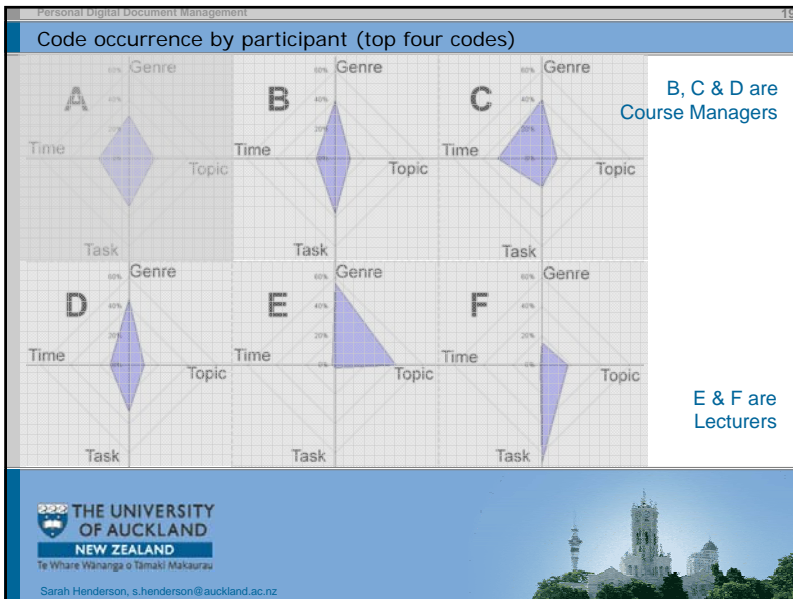
Proportion of folders with each naming code

Code Descriptions

Code	Description & Examples
Genre	Indicates that the contents of the folder are a particular class or type of document, with a commonly recognized form and structure. Examples: Lecture Notes, Presentations, Timesheets, Budgets, Letters.
Task	Indicates that the contents of the folder are related to a task, project, event or some other type of activity. Examples: Assignment 5, Lec01, PhD, recruitment, For DSS Presentation.
Course	Indicates that the contents of the folder are related to a specific course. (This is a special case of Task above) Examples: Database Systems, 222, INFOSYS 222
Topic	Indicates that the contents of the folder are all about a particular subject matter. Examples: Web development, Database Architectures, JavaScript

---

Code occurrence by participant (top four codes)

B, C & D are Course Managers

E & F are Lecturers

Folder hierarchy

Participant		Scheme	Confidence
	A	Time > [various]	Low
Course Managers	B	Time > Course > Task	Medium
	C	Genre > Time	Medium
	D	Task > Course > Time > Genre	High
Lecturers	E	Task > Time > Course > Genre or Task > Course > Time > Genre	High
	F	Genre/ Task > [various]	Low

Facets

```
⊟ 📁 Labs                    ⊟ 📁 INFOSYS 221          ⊟ 📁 2004
  ⊟ 📁 2004                    ⊟ 📁 2004                 ⊟ 📁 INFOSYS 221
    📁 INFOSYS 221                📁 Labs                    📁 Labs
    📁 INFOSYS 222                📁 Lectures                📁 Lectures
  ⊟ 📁 2005                      📁 Tutorials               📁 Tutorials
    📁 INFOSYS 221             ⊟ 📁 2005                 ⊟ 📁 INFOSYS 222
    📁 INFOSYS 222                📁 Labs                    📁 Labs
⊟ 📁 Lectures                    📁 Lectures                📁 Lectures
  ⊟ 📁 2004                      📁 Tutorials               📁 Tutorials
    📁 INFOSYS 221           ⊟ 📁 INFOSYS 222          ⊟ 📁 2005
    📁 INFOSYS 222            ⊟ 📁 2004                 ⊟ 📁 INFOSYS 221
  ⊟ 📁 2005                      📁 Labs                    📁 Labs
    📁 INFOSYS 221                📁 Lectures                📁 Lectures
    📁 INFOSYS 222                📁 Tutorials               📁 Tutorials
⊟ 📁 Tutorials               ⊟ 📁 2005                 ⊟ 📁 INFOSYS 222
  ⊟ 📁 2004                      📁 Labs                    📁 Labs
    📁 INFOSYS 221                📁 Lectures                📁 Lectures
    📁 INFOSYS 222                📁 Tutorials               📁 Tutorials
  ⊟ 📁 2005
    📁 INFOSYS 221
    📁 INFOSYS 222
```

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

Desktop Screenshots

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

File Management vs Document Management

- Distinction between files and documents

- Example: a status report that went through five drafts, was edited once by the boss and sent to a client.
- This is actually six separate files in the file system plus two in the email system, with no relationship between any of them
  – except perhaps a similar file name, but that is up to the user

- An interface that recognizes and manages documents (rather than files) could help overcome the version management problems reported by these participants.

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

Useful Distinctions between users

• Browse oriented	• Search oriented
• Post-structure	• Pre-structure
• Deleter	• Hoarder
• Conceal oriented	• Display oriented
• Tree averse	• Tree oriented
• Subject/Project oriented	• Time oriented
• Specify names	• Use default names

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Sarah Henderson, s.henderson@auckland.ac.nz

Conclusion

- There is no "best" way to combine these dimensions into a hierarchy.
  - Forcing them into a hierarchy results in duplication
  - A facet-based document management system would be a promising approach

- Some opportunities for automatic software support of document management (Person, Source, Topic, Time, File Type), but more research needed.

- More research needed on genre in personal digital document management

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
Sarah Henderson, s.henderson@auckland.ac.nz

Questions?

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
Sarah Henderson, s.henderson@auckland.ac.nz

References

- Dourish, P. et al. (1999) Presto: An Experimental Architecture for Fluid Interactive Document Spaces. ACM Transactions on Computer-Human Interaction 6 (2), 133-161

- Freeman, E. and Gelernter, D. (1996) Lifestreams: A Storage Model for Personal Data. SIGMOD Bulletin 25 (1), 80-86

- Rekimoto, J., Time Machine Computing: A time-centric approach for the information environment. in UIST'99 Symposium on User Interface Software and Technology, (Asheville, North Carolina, USA, 1999), 45-54.

- Robertson, G.G. et al. (1998) Data Mountain: Using Spatial Memory for Document Management. In UIST'98 Symposium on User Interface Software and Technology, pp. 153-162, ACM Press

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau
Sarah Henderson, s.henderson@auckland.ac.nz