

Preserving 20th & 21st Century Digital Lives: MyLifeBits, a “digital life” experience

10 February 2009

British Library First Digital Lives Conference

Gordon Bell research.microsoft.com/~gbell

Microsoft Research

Founder and Trustee Computer History Museum

The Digital Lives Project

- The project is considering not only how collections currently being deposited are changing
- **Establishing how modern personal digital collections are being created, managed, and made accessible (either now or in the future)**
- Exploring the needs and views of potential scholarly users of future personal digital collections such as biographers and historians
- Establishing how curators should approach selection, description, preservation and access for these personal digital collections
- **Identifying ... the implications and methods of transfer of personal digital collections to long-term repositories**
- Researching and assessing potentially transferable tools ...
- Establishing the impacts of legislation, confidentiality and ... ethics ...
- Identifying how we should address “hybrid” personal collections of ...
- Researching new services ... for managing and publishing personal digital collections and assessing their potential effectiveness for preservation
- Sharing our experience and outcomes from this research with others

A View of Preserving Digital Lives

- Preserving the analog life of a 20th century person:
10-100s GB. 2 Mpages, 100 Kimages, 100-1,000 hrs. video
 - Won't analog people need to be converted to digital,
... if not they're really gone and forgotten history?
 - Gresham's Law: digital lives may drive out analog lives
- How will a 21st century, digital person be preserved?
 - Which "lives" of a person e.g. personal, professional?
 - Depth of each life?
 - Size. Who's in the BL's digital lifeboat?
 - Can 20th archivist track evolving "digital" people?
- Half-breeds seeking sanctuary arriving e.g. Salman Rushdie
 - Notes, manuscripts, books, videos, etc.
 - 5 MACbooks and 60 GB hard drive
- Preserving Everybody?
 - Role of public institutions vs. the cloud for "all of us"

Digital Archivists In Demand NY Times 7 February

- WHEN the world entered the digital age, a great majority of human historical records did not immediately make the trip.



» As preservation officer at U.C.L.A., Jacob Nadal safeguards materials (digital and analog) in its collection

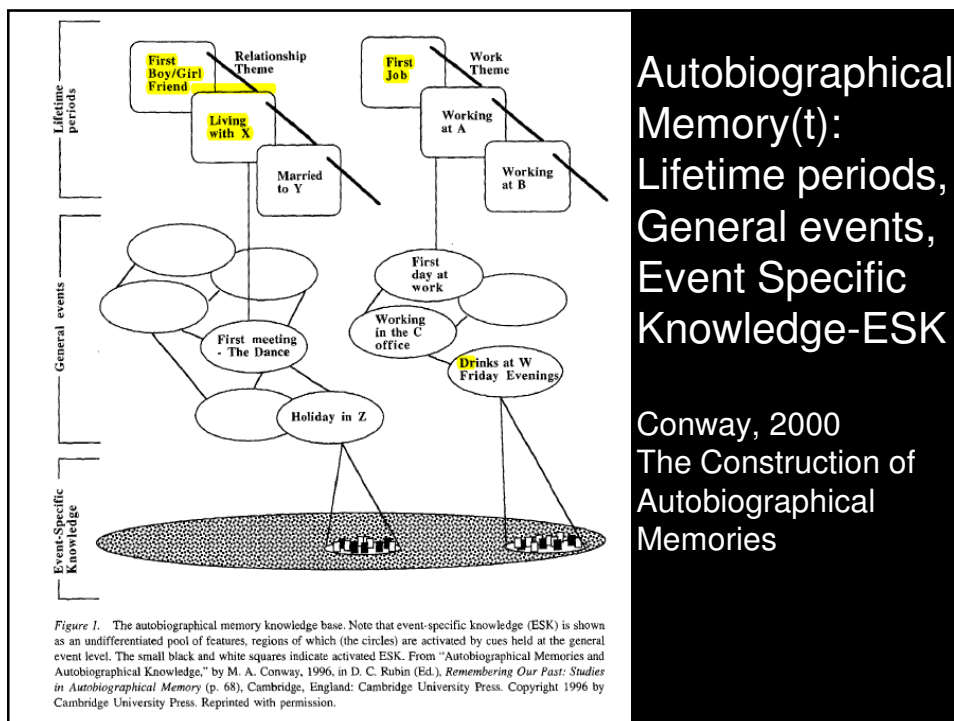
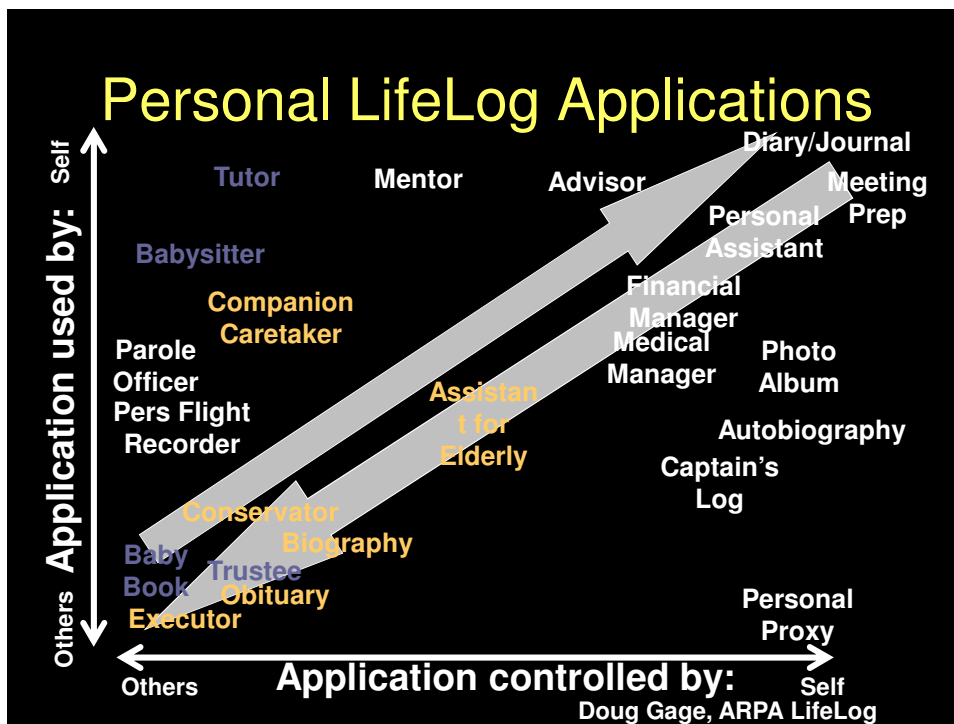
- Literature, film, scientific journals, newspapers, court records, corporate documents and other material, accumulated over centuries, needed to be adapted for computer databases. Once there, it had to be arranged — along with newer, born-digital material — in a way that would let people find what they needed and keep finding it well into the future.

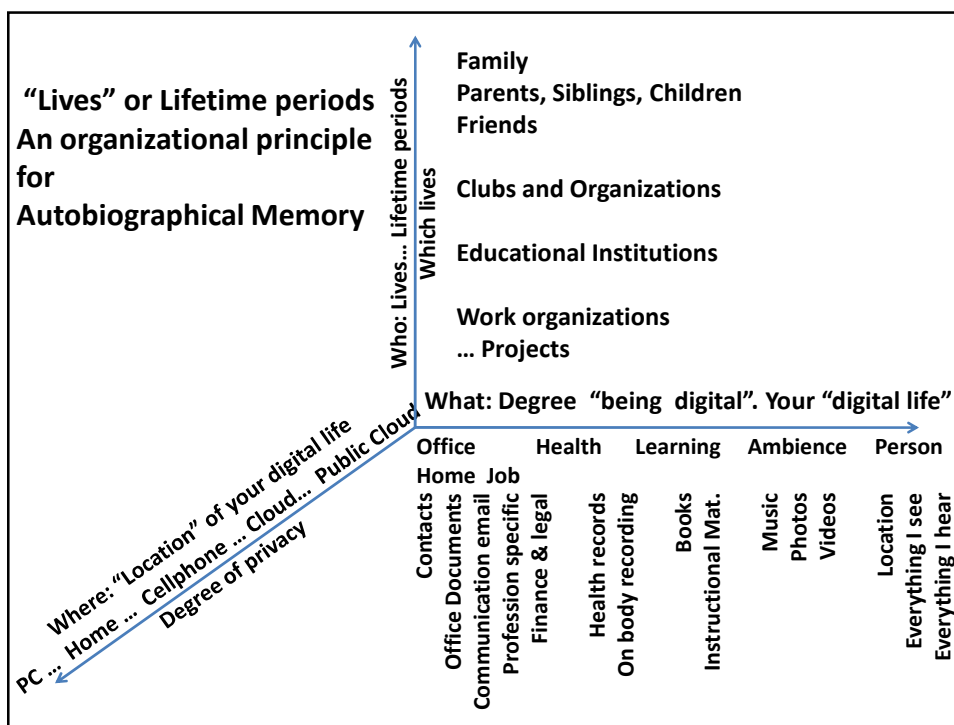
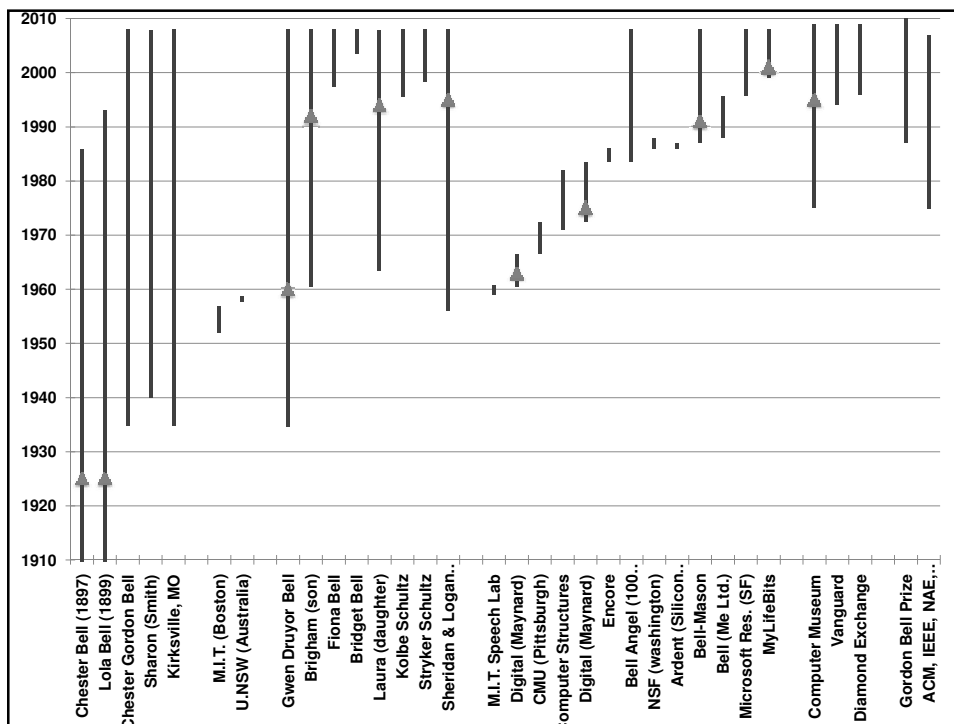
GB Outline

- **I. MyLifeBits: Experience with an early digital life**
 - **Why** (NOT a narcissistic exercise. Evolved from being telepresence, ambiance and paperlessness, triggered by Raj Reddy Million book project. Bush, and Gates. In 1998: curiosity about size, shape, cost, in light of a Terabyte drive. An inevitable quest for the PC ... but not the mind/memory, PRODUCT, and never ending research. Now it is about the mind & immortality!
 - **How much of my life was recorded?** >9/10 mundane & repetitive. Conversations nice.
- **II. Implications for digital “life support”**
 - **Who. Which “digital lives” or Lifetime periods?** (family, friends, youth, schools, organizations, sports, churches, companies,... projects)
 - **What. General events. Life context?** (Memory (to remember and remind, provide a record); office at home, work; learning; health; ambiance, fun and entertainment; immortality)
 - **What. Event Specific. Which of life’s activities? *Depth*** (scans, born digital, docs, communication, books, photos, videos, songs, records, reminders,...SenseCam, meetings and phone calls, GPS trajectories (x, y, z, t, r, phi, theta) and where looking, wellness activities, health records, ... phone)
 - **Where is your digital life and how public is it?** PC, cloud, cell phone, svc
 - **How many lives: Is Every *digital* Life Worth Saving?** Yes, it is possible.
 - **How do we make sense of a life or of multiple lives?** Very interesting!
- **Questions *that are often of interest to ... lives, institutions***
 - **Appy, “...will you still support me when I’m (you’re) old?” Signed, Lost Data**
This is the most often sighted problem for archivists in light of the advantages of paper.
 - **Where is my digital life?**
 - **How deep is the cyber landfill i.e. personas life? Map of where to dig to find?**
 - **Will my data be: private, secure, available to me, taken *if I don’t own or holdit***

MyLifeBits Manifesto

- An individual’s digital life (information) is accumulating rapidly
- Digital information cost to store, transmit and replicate is negligible
- The information about an individual, including what one heard, said, and saw, has a super-linear value that *exceeds acquisition & retention cost*.
 - **to supplement, support, and augment bio-memory**
 - **to free one’s life from clutter i.e. reduce space & energy**
 - **to assist in other ways e.g. ambiance, communication, dematerialization, telepresence (work anywhere)**
 - **to enable digital immortality**
- Individuals must take on the responsibility for recording, managing and retaining the increasing number of evolving digital information streams i.e. computers, phones, IPTV, health sensors and legacy physical bits.
- **Nothing should be deleted**
 - A unified store for anywhere, anytime access while satisfying privacy, sharing, management, and retention is essential.





MyLifeBits organization: time and space (lives)

**Lives/
Context
(space)**

**Personal
(some \$s)**

**GB Co.
(angel, etc.)**

**Professional
ACM, etc., ...
@Microsoft.com,
New co's.**

Archival (time)

Working

Financial,
Legal,
Health, etc.
Records



MyLifeBits Bits

Everything goes in a database

- MyLifeBits needs all the features of a database (Consistency, Indexing, Pivoting, etc., Speed/scalability, Backup, Replication)
- If we didn't use a database, we'd have to manually create one!
- Files as blobs; search engines for legacy apps
- We were part of the Microsoft Research Area Research Lab



SQL

Family History (Relics of the Past)



Family Mementos and Pictures

Family Records and Documents

Personal History (Relics of the Future)



Artwork, Collections, Unique
Personal Story Mementos


Potential Family History
Documents, Pictures,
Home Movies

Professional Life (Historical Record)



Awards, Honors, Certificates, Medals
Notebooks, Calendars, Books,
Reports and Papers, Video

Professional Life Office



Stored papers, calendars, and notebooks

Recording Everything:

10/98: Raj Reddy request to scan books

12/00: 19.8 GB, 12 K items, 100 Kpg, 5K office docs, mail, books, photos, CDs, 20 Videos.

Items types: bills, brochures, cards (bus., greeting ... report), certificates, clippings, ... manuals, memorabilia, receipts

01/01: Challenge is not recording everything. It is getting it back.

10/01: Decide to use SQL

12/02: FIRST Paperless year. 1' scans, 12' trash.

TV and Radio programs,

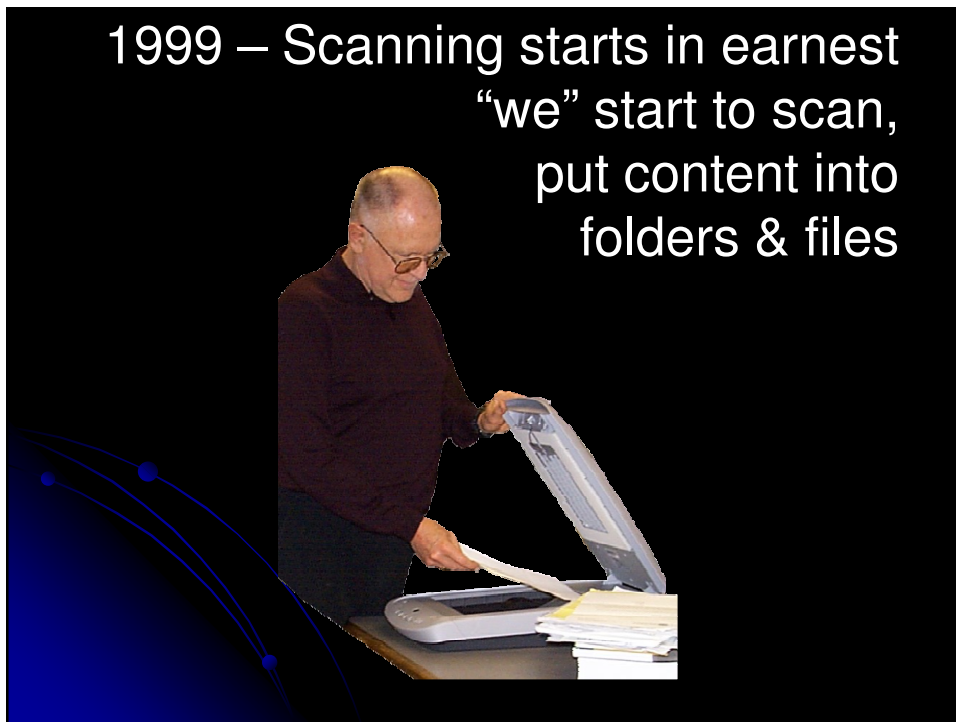
02/03: Techfest MyLifeBits demo with relational database; text & voice comments, every web page, timelines, ISBQ

10/03: SenseCam (Everything seen), Vibe (Keybrd, mouse clicks)

08/04:GPS tagging of photos for trip replay

05/06: MyLifeBits is a personal TP. Phone calls, rooms, meetings

1999 – Scanning starts in earnest
“we” start to scan,
put content into
folders & files



1999 – How did you do all that
scanning?

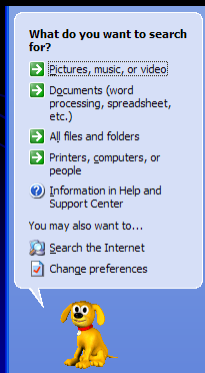


Now that it's in Cyberspace

How do you remember the 30,000+ file names?

Or in which of 2500 folders they live?

What's about a tool for finding stuff?



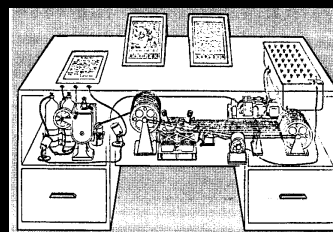
Re-discovery of Memex

As We May Think, Vannevar Bush, 1945

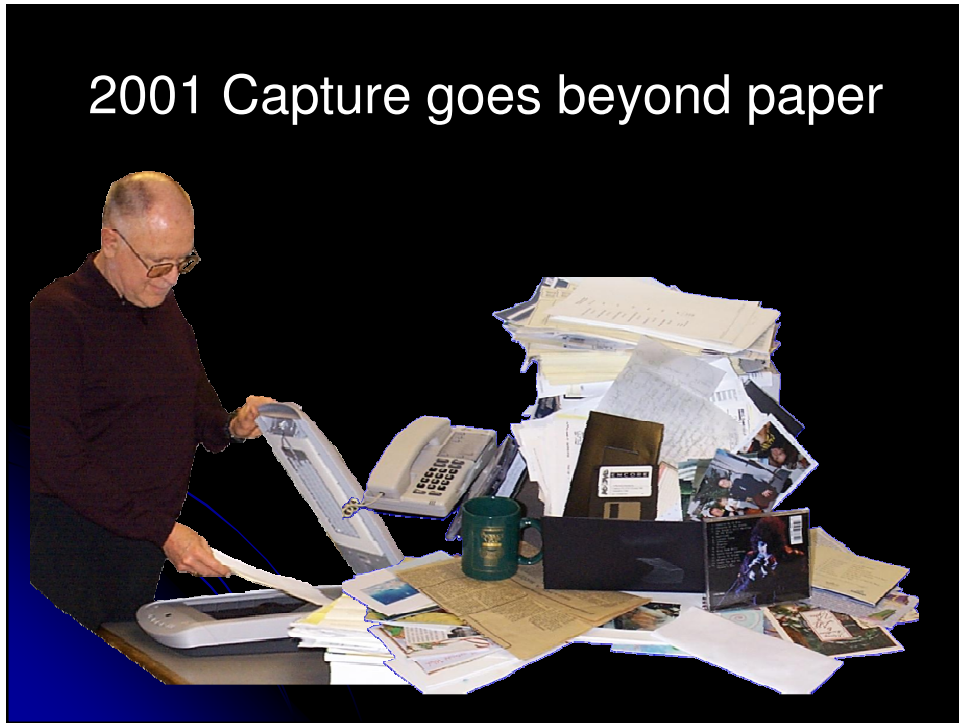


“A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility”

- Full-text search, text & audio annotations, and hyperlinks

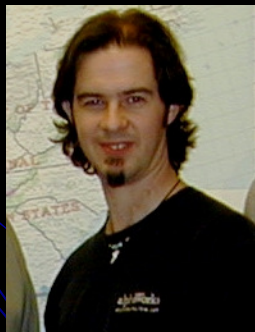


2001 Capture goes beyond paper



Oct 2001

- Gemmell & Lueder start building system



Everything goes in a database

- MyLifeBits needs all the tricks of a database (Consistency, Indexing, Pivoting, Replication, Speed/scalability, Backup, Replication)
- If we didn't use a database, we'd have to manually create one!
- Files as blobs; search engines for legacy apps
- We are part of J... Area Research Lab

Search for text: brig

Type

- Annotations
- Attachments (32)
- Contacts (4)
- Outlook types (1171)
- Communications (11135)
 - Appointments (16)
 - E-Mail (1115)
 - Notes
 - Phone Calls
 - Tasks (4)
- Documents (385)
 - Presentations (10)
 - Scans (19)
 - Spreadsheets (38)
 - Text Documents (80)
 - Web Pages (238)
- Files (1733)
- Folders (18)
- Music (2)
- Pictures & Video
 - Pictures (1567)
 - Radio shows
 - TV Shows
 - Videos (22)

Date

Type Pictures

Path

Rating

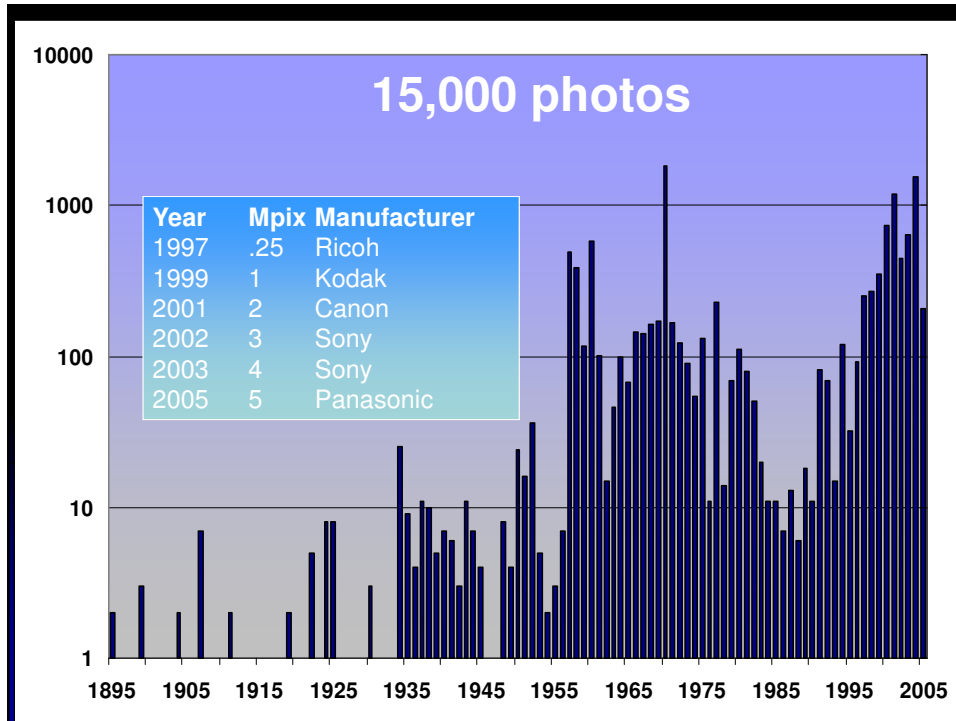
Location

Camera Make

Camera Model

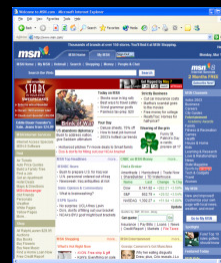
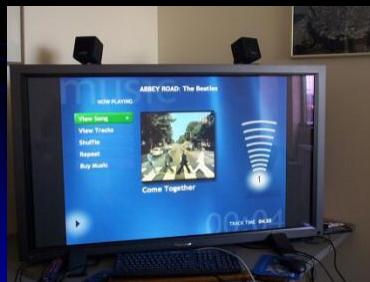
1961

Jan	Feb	Mar	Apr	May	Jun	Jul

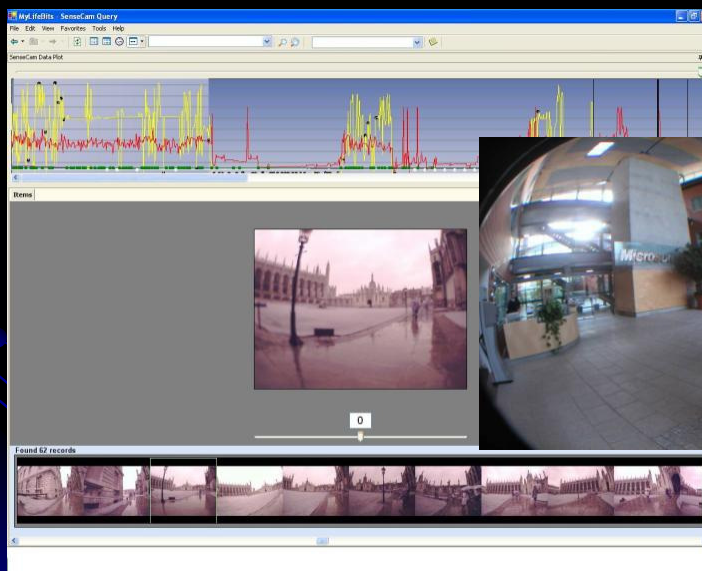


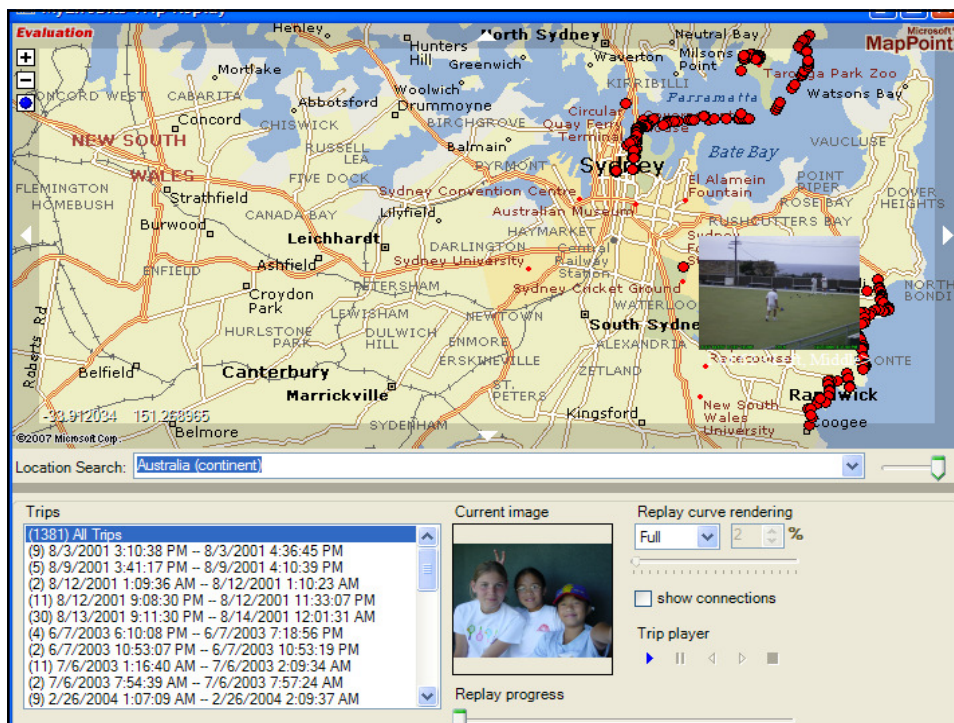
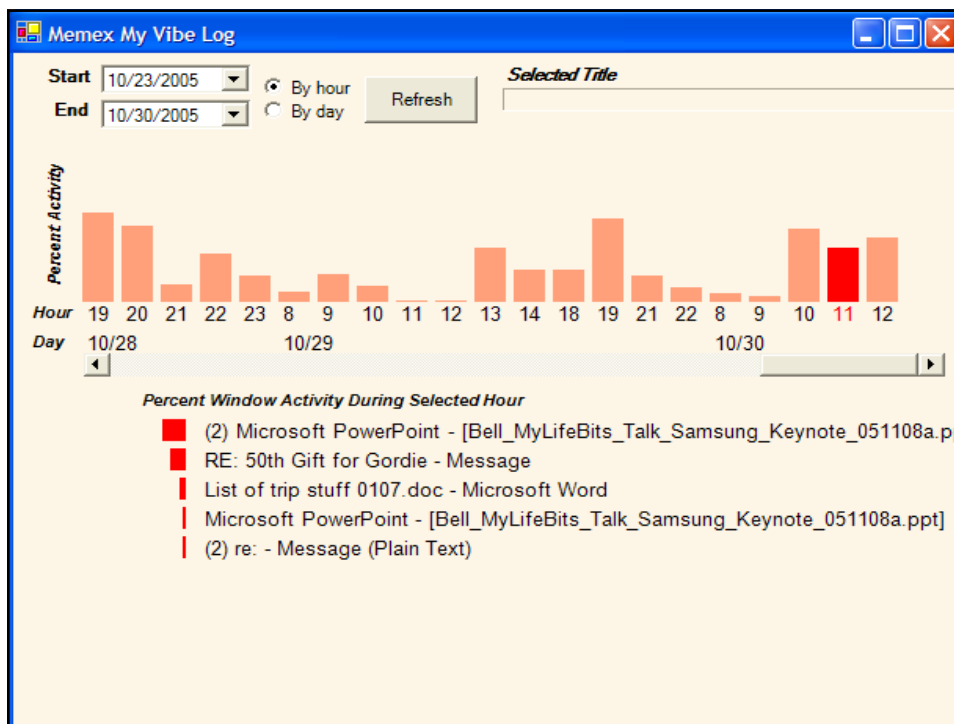
Even more capture

- Telephone calls, more video, all web pages visited, usage logging, radio, TV...



2003 - SenseCam



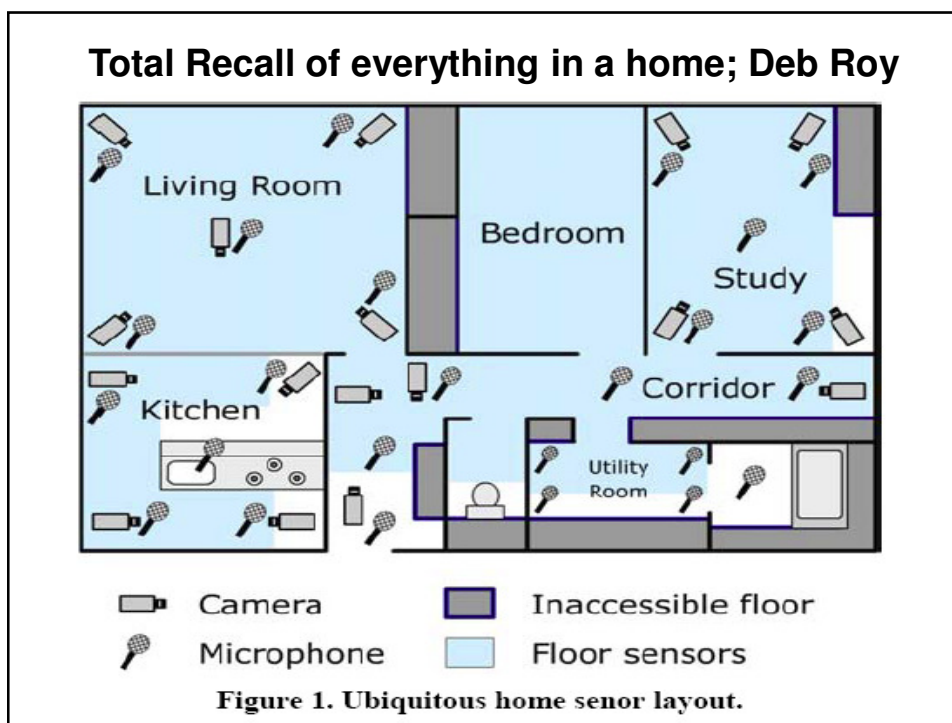
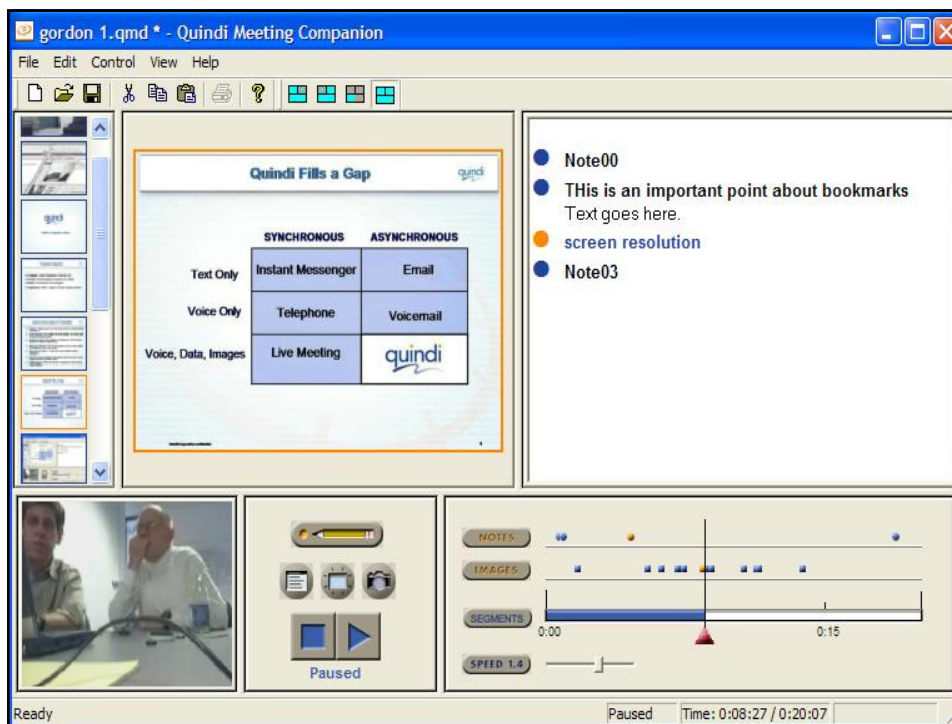


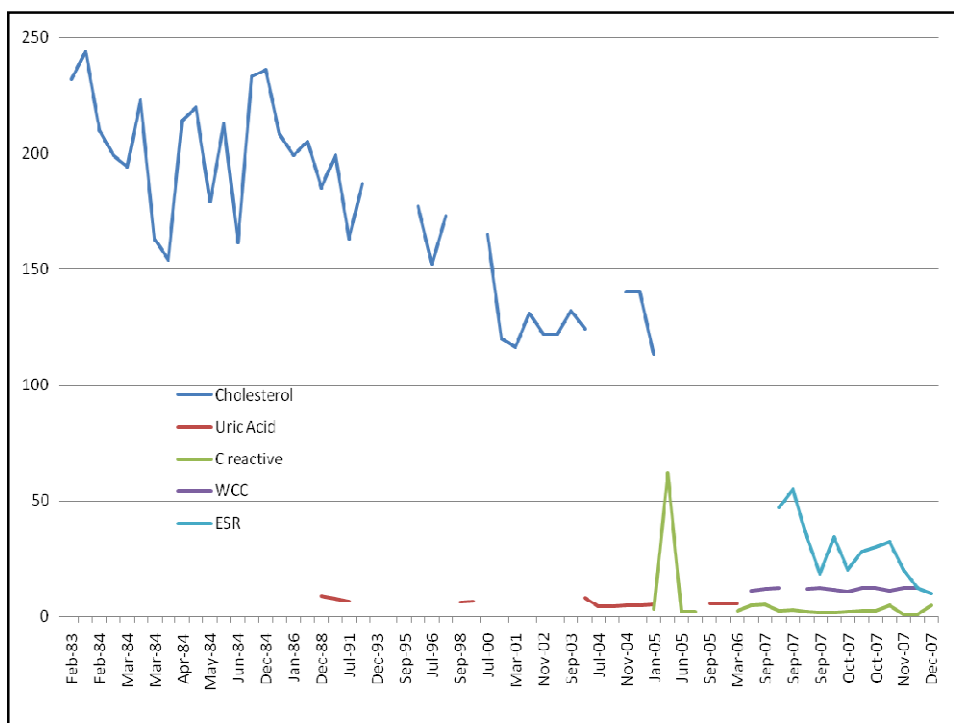


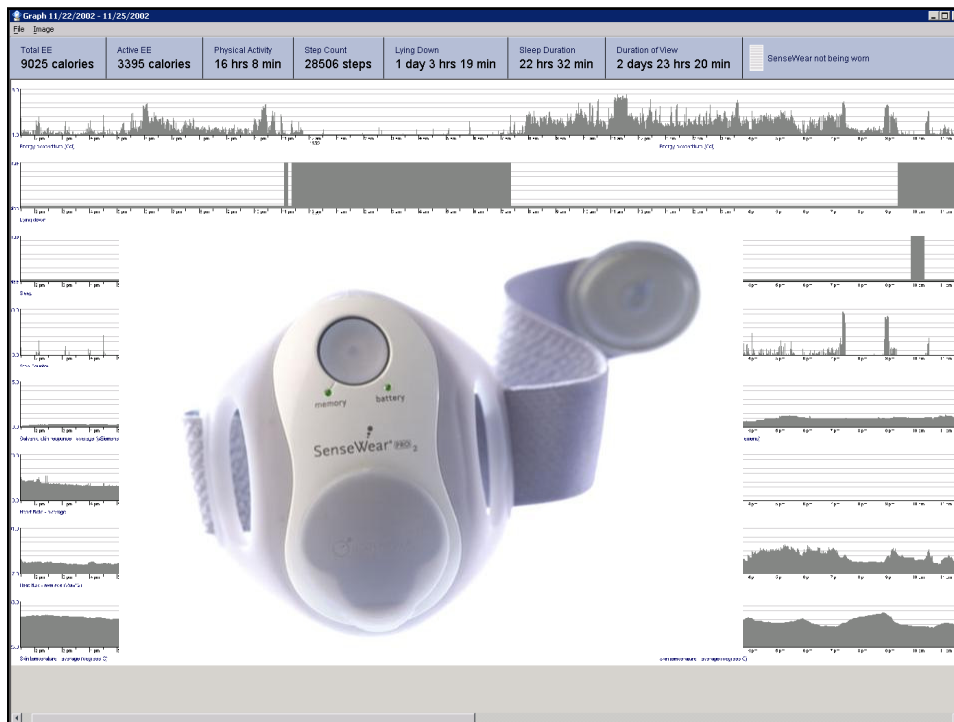
Feb 2005

Epiphany!

Memex is a database
&
personal TP system





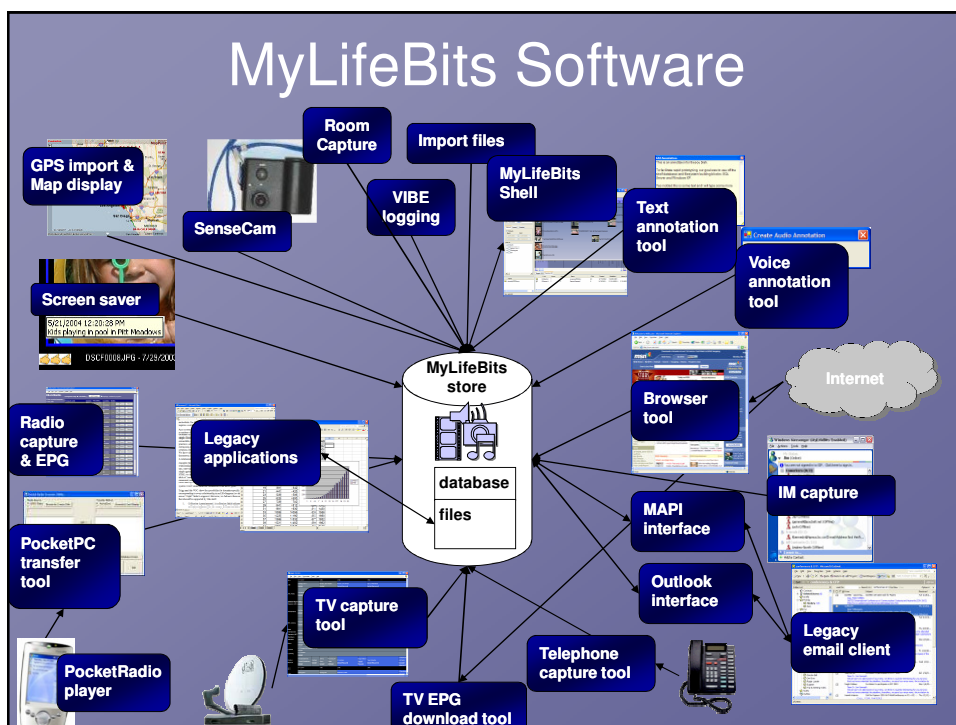


Capturing every heartbeat

- 72.6 beats/min; 38.16 Million beats/year
- 3.13 billion beats year
- The important number is 4-4.5 years, or ETS
- Battery life: the expected trip to surgery!

Rate (ppm)	Event Counts				
	PV	PR	AV	AR	PVE
30 - 54	411,860	4	28,630	1	0
55 - 69	6,824,410	195	4,614	3,609	12
70 - 89	8,113,024	1,359	0	0	274
90 - 109	2,516,074	524	0	0	386
110 - 129	451,814	212	0	0	180
130 - 149	12,599	104	0	0	114
150 - 179	292	46	0	0	112
180 - 224	0	1	0	0	6
225 - 249	0	2	0	0	2
> 250	0	0	0	0	3
Total:	18,330,073	2,447	33,244	3,610	1,089
	Total Event Count:		18,370,463		

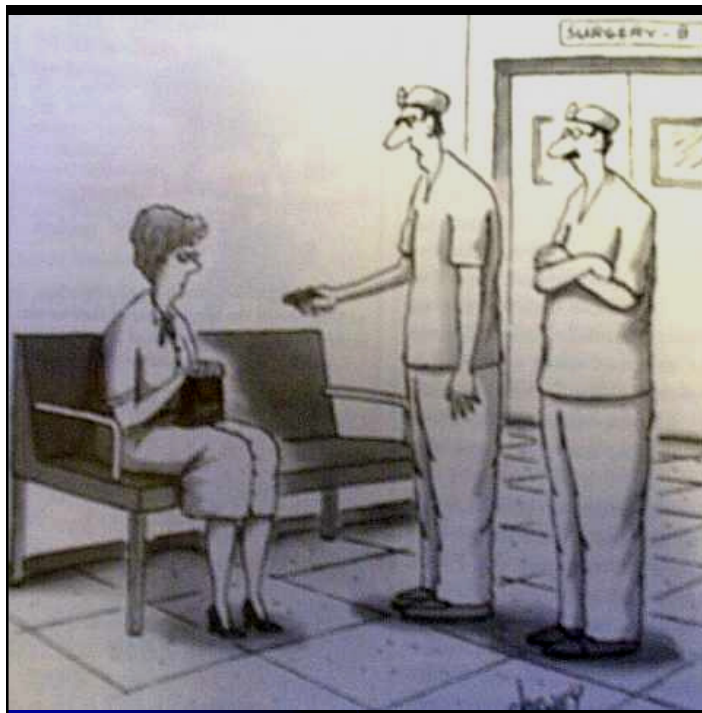
The “killer app”... Health?



What do you carry with you and in you?

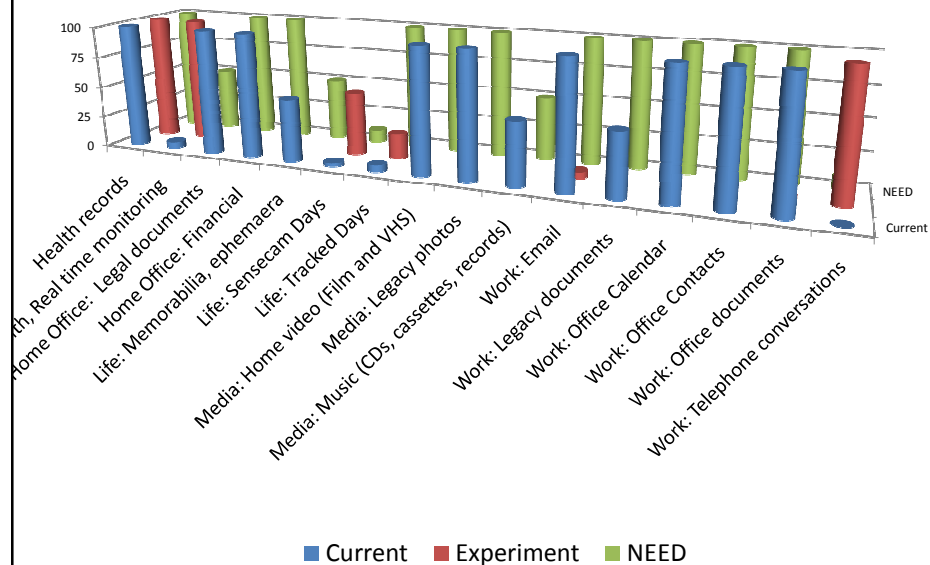
1. Laptop & PS (main computer) ... backed up in the cloud
2. Cell phone & PS
3. Ipod & PS (in cable carrier)... 8,000 songs!
4. Earbuds: small, no power
5. Kindle & PS (many books)
6. Kodak & PS (camera)...normally carry two
7. GPS tracker (PC or USB charger)
8. Sensecam; Audio recorder & batteries
9. BodyBugg (batteries) for energy, HR, monitor
10. **Implanted** Pacemaker 5 year battery life (**not shown**)
11. Cable carrier (techno stuff--connect anything anywhere)
12. *Spare laptop & PS (spare)*
13. Razor & PS
14. ~~Bose Case w/ PS very bulky noise cancelling earphones~~
15. ~~Back up hard drive...~~

Smart
Phone



Your
husband
just died,
... here's
his black
box

How much of life need be recorded



Content type (5)

- Audio
 - Clip
 - Man made
 - Industrial
 - Sound effects
 - Natural
 - Bird
 - Water
 - Running
 - Waves
 - Song
 - Voice
 - Answering machine
 - Comment or annotation
 - Conversation
 - Meeting
 - Memo or dictation
 - Presentation
 - Speech
 - Telephone cal
- Document type (180)
- Graphics
 - Art
 - Album art
 - Cartoons
 - Drawings
 - Paintings
 - Diagrams
 - Graphs
 - Maps
 - Schematics
- Photograph
 - People
 - Group
 - Portrait
 - Scenery
 - City
 - Domestic
 - Rural
- Video
 - Commercial
 - Home

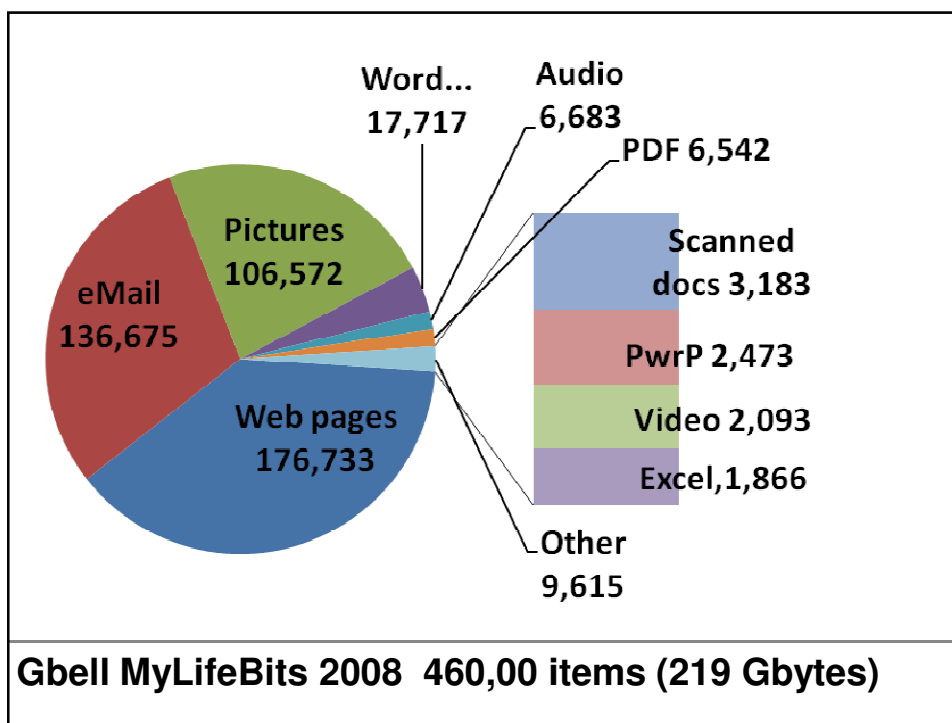
Content type, document type

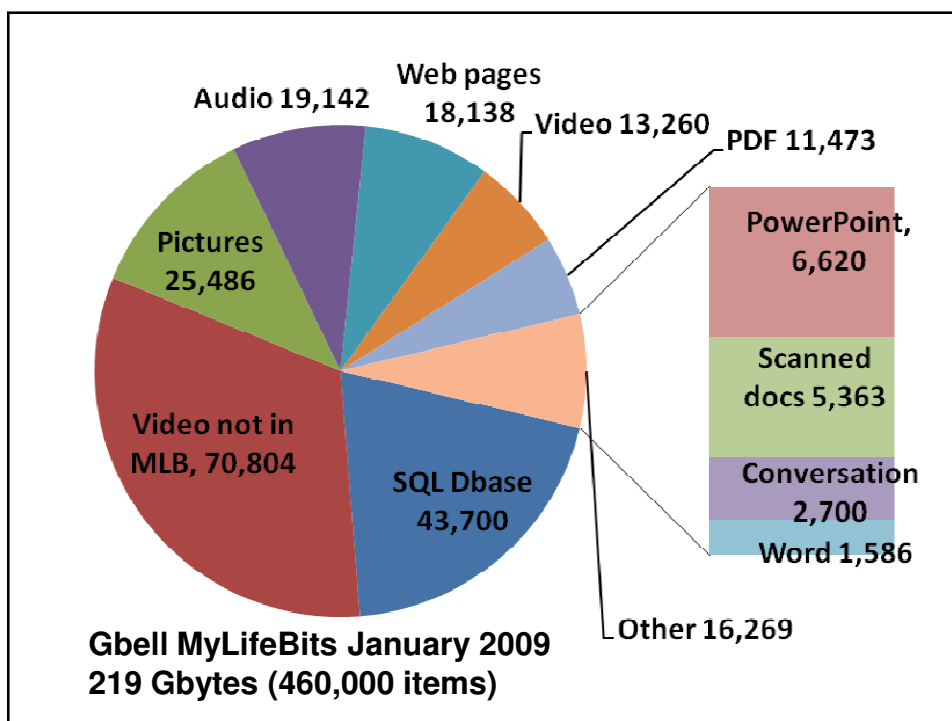
Document type (180)

- Advertisement (5)
- Ballot
- Brochure
- Business card
- Calendar
- Certificate (14) achievement, birth...diploma,...license...stock, warrantee
- Commentary @ analysis (4) annotation, essay, paper rev., performance rev.
- Computer source code
- Correspondence (4) announcement, invitation, newsletter, recommendation
- Description (1) Standard specification
- Eulogy
- Exam
- Financial
 - Bill, statement, invoice (3)
 - Budget
 - Estimate
 - Financial report
 - Annual report
 - Expense report
 - Purchase order
 - Term sheet
- Form (6) application, COA, questionnaire, renewal, survey, tax form & return
- Grade card
- Greeting card
- List (15) bibliography, computer program output, database report, experimental results, grocery, inventory, laundry, menu, price, references, roster, service, statistics, web
- Name tag
- Newsgroup posting
- Obituary
- Plans (4) business, goals, proposal, schedule ((3) agenda, order of service, program)
- Poetry
- Presentation
- Record of events (5) biography & auto biography, diary, news, progress report
- Reference work dictionary, encyclopedia, web reference
- Reservation
- Resume
- Rules and guidelines
- Scrapbook
- Script (dramatic)
- Summary & syllabus
- Ticket or stub
- Transcript (computer chat, interview, speech)

The screenshot shows a hierarchical tree view of subject categories on the left and a list of organization types on the right. The left pane is titled 'subject' and includes categories like Career, Computers, Hardware, Software, Education, Family, Finance + Investing, Financial, Personal Property, Home, Shopping, Travel, and Underwater. The right pane is titled 'Collections' and shows a list of organization types such as Dissemination (format, who), Event Types, Institution, and Organization (Company, non-profit, government agency). Below these are various specific organizations like Ardent, Churchill club, Cirrus, Computer Museum, Dana, DES, Digital, and Microsoft.

Subject (old), Organization, Organization type





MyLifeBits Manifesto

- An individual's digital life (information) is accumulating rapidly
- Digital information cost to store, transmit and replicate is negligible
- The information about an individual, including what one heard, said, and saw, has a super-linear value that *exceeds acquisition & retention cost*.
 - **to supplement, support, and augment h-memory**
 - **to free one's life from clutter i.e. reduce space & energy**
 - **to enable digital immortality**
 - **to assist in other ways e.g. ambience, communication, dematerialization, telepresence (work anywhere)**
- Individuals must take on the responsibility for recording, managing and retaining the increasing number of evolving digital information streams i.e. computers, phones, IPTV, health sensors and legacy physical bits.
- **Nothing should be deleted**
- A unified store for anywhere, anytime access while satisfying privacy, sharing, management, and retention is essential.

Personal Archive Examples

- Charles Vest: former President of MIT
- Einstein
- Allen Newell: CMU
- Lederberg: National Lib. Of Medicine
- Comparisons with MLB

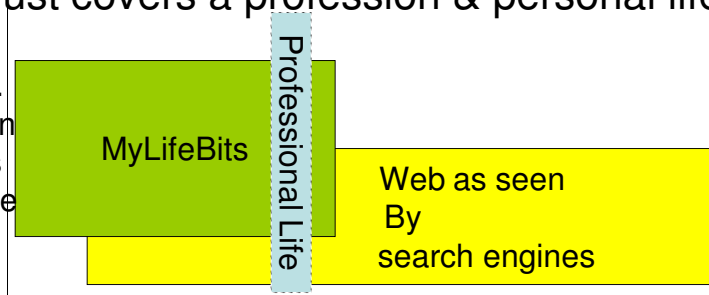
Public 21st century figure legacies

- Charles Vest, president of MIT from 1990 to 2004, delivered a hard drive with nearly all of the files of his 14 year tenure to the MIT Archivist. It including speeches and letters (drafts), presentations, planning documents, meeting minutes, e-mails and a few photos. The only items Vest had deleted were a few files about his personal finances.
- Nothing had been scanned, so no incoming input such as letters, web page views, articles, unless they were attachments.

Personal Search is not Professional or Web search

- System sees every entry & access
- Everything, not just a professional life
- Life is limited to SIS (stuff I've seen), not all of it. Just covers a profession & personal life

Depth e.g.
information
item types
& coverage



Knowledge breadth e.g. Dewey classification

Einstein Archives Online

Albert Einstein

- Digitized Manuscripts
- Finding Aid
- Archival Database
 - Itemized database of approximately 43,000 Einstein and Einstein-related archival items: writings, professional & personal correspondence
- Gallery
- About
- Terms of Use
- Credits
- Feedback

A joint project of the

Albert Einstein Archives
and the David and Fela Shapell Digitization Project
at the Jewish National & University Library,
the Hebrew University of Jerusalem
האוניברסיטה העברית בירושלים

Einstein Papers Project
at the California Institute of Technology
in Pasadena

Last updated: May 19, 2003*

www.alberteinstein.info

Einstein Archives Online

Digitized Manuscripts
 Finding Aid
 Archival Database

A. Scientific Material

- A.1 Scientific Manuscripts and Notebooks**
 - A.1.1 Scientific Manuscripts**
 - A.1.2 Scientific Notebooks**

B. Non-Scientific Material

- B.1 Non-Scientific Manuscripts**
 - B.1.1 Non-Scientific Holographic Manuscripts, arranged chronologically**
 - B.1.2 Typescripts of Non-Scientific Manuscripts, arranged chronologically**
 - B.1.3 Other Non-Scientific Manuscripts, arranged chronologically**

C. Biographical Material

Libraries/UnivArchives: Allen Newell Collection [Carnegie Mellon Libraries] - Microsoft Internet Explorer

Address <http://diva.library.cmu.edu/newell>
Go

Welcome to the full-text digital archive of Allen Newell's pioneering work in artificial intelligence and cognitive science. Newell earned an international reputation as one of the founders of artificial intelligence. He is also known for his work in the theory of human cognition and the development of computer software and hardware systems for complex information processing.

[Access](#) → Access the Newell Collection containing publications, project reports, teaching materials, correspondence, and eventually audio and video.

[Contents](#) → Explore the scope and content notes of the Newell Collection.

[Biography](#) → Learn about Allen Newell's background, those who influenced his life and work, his collaboration with Herbert A. Simon, and other interesting information.

[Exhibit](#) → The roots of artificial intelligence (AI) are revealed. The pioneering work of professors Herbert A. Simon and Allen Newell inspired advancements in cognitive theory and computer science worldwide.

[Resources](#) → Visit other sites with tributes, bibliographies, and related information.

Gabrielle V. Michalek, Head of Archives Digital Library Initiatives
 Web site design by [Maisy Hickey](#), Computer Science Librarian
 November 14, 2003 -- <http://diva.library.cmu.edu/Newell/>

Libraries/UnivArchives: Contents [Carnegie Mellon Libraries] - Microsoft Internet Explorer


Address: <https://diva.library.cmu.edu/Newell/contents.html>

Contents

Allen Newell Collection


The Allen Newell Collection comprises 116 cu. ft. of papers arranged in 12 topical series. The collection also contains 12 cu. ft. of audio cassettes, 35mm slides, videotapes, photographic negatives, glass plate (lantern) slides, and computer tapes.

The collection includes: scientific papers, project reports, manuals and proposals by Newell, his Carnegie Mellon colleagues and others; computer programming instructions, printouts, and cognitive experiment data; teaching materials and student files containing papers and correspondence; lecture materials and conference information; publications and journal reprints; and external correspondence, Carnegie Mellon interoffice memoranda, and e-mail printouts.



Series Comprising the Collection

Series I: RAND (The RAND Corporation)	Series VII: Conferences and Professional Organizations
Series II: Soar	Series VIII: ARPA
Series III: Human Problem Solving	Series IX: Publications
Series IV: William James Lectures	Series X: Programming Instructions
Series V: Unified Theories of Cognition	Series XI: Correspondence
Series VI: COGNET	Series XII: Computer Science Department



A working knowledge of artificial intelligence, computer science, and cognitive psychology would assist in understanding the collection. Recommended reading includes *Human Problem Solving* and *Unified Theories of Cognition*, the Newell memorial items contained in the Publications series, as well as Herbert Simon's autobiography *Models of My Life* (1991), pp. 135-168, 189-234, etc.

National Library of Medicine

- Top 30: 265 GB; 181K Files; 1.5 MB/file
 - 99.99% tiff images
 - less than .01% plain text
 - less than .01% html files
 - less than .001% AVI video
- Web derivative files 36 GB; 75K files; 0.5 MB/file
- Items Pages Video Who

18,615	49,951	8	Lederberg
1,738	37,110	25	RMP
1,054	10,811	5	Koop
580	4,374	1	Avery
469	1,833	-	Crick
279	1160	1	Pauling
302	893	-	Varmus
...			
27.5 K	143 K		


Profiles in Science: The Joshua Lederberg Papers - Microsoft Internet Explorer

Address: http://profiles.nlm.nih.gov/BB/ lederberg

Profiles in Science National Library of Medicine

Home | Collection Home | Search | Browse | What's New | About

The Joshua Lederberg Papers



- Biographical Information
- Bacterial & Cellular Genetics and the Nobel Prize
- Artificial Intelligence and Expert Systems
- Exobiology
- Emerging Infectious Disease and Biological Warfare
- Health and the Future

[All Documents](#) [All Visuals](#)

Joshua Lederberg (b. 1925) is an American geneticist and microbiologist who received the Nobel prize in 1958 for his work in bacterial genetics. The National Library of Medicine is the repository for the Joshua Lederberg Papers, which range from 1910 to 1990s, the majority dating from 1945 to the present day. The collection contains laboratory notebooks, research reports, published articles and books, correspondence, unpublished manuscripts, speeches, news clippings, photographic prints, and audiovisual materials.

As part of its Profiles in Science project, the National Library of Medicine is digitizing and making available over the World Wide Web a selection of the Joshua Lederberg Papers, for use by educators and researchers. This Web site provides access to the portions of the Joshua Lederberg Papers that are now publicly available. Individuals interested in conducting research in the Joshua Lederberg Papers are invited to [contact](#) the National Library of Medicine.

This online Exhibit is designed to introduce you to the various phases of Dr. Lederberg's scientific career and professional life. It is divided into sections that focus on Lederberg's life and major scientific contributions. We suggest that new visitors begin with this exhibit, which includes a small selection of documents and visuals, organized within these sections. Each section begins with a "Background Narrative," which leads to "Documents" and "Visuals."

Visitors may access additional materials through [Search](#) on the navigation bar. They may also view the materials alphabetically or chronologically by choosing [Browse](#) on the navigation bar. Documents and visuals in these lists are arranged by format and then either alphabetically by title or chronologically.

U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health, Department of Health & Human Services
[Copyright](#) [Privacy](#) [Accessibility](#)
[Comments](#) [Licenses](#) [Acknowledgments](#)



























































The Joshua Lederberg Papers: Official reports (Chronological Listing) - Microsoft Internet Explorer

Address: http://profiles.nlm.nih.gov/BB/Views/AlphaChron/date/10039/ lederberg

The Joshua Lederberg Papers


Home | Collection Home | Search | Browse | What's New | About

Official reports (Chronological Listing)

- [Excerpts from "Economic Benefits from Oceanographic Research"] (1964)  
- [ACME Annual Report, 1968] (1968)   
- [ACME Annual Report, 1969] (1969)   
- [ACME Annual Report, 1970] (1970)   
- [ACME Annual Report, 1971] (1971)   
- "Resource-Related Research - Computers and Chemistry" Annual Report, 1972 (September 30, 1972)   
- [ACME Annual Report, 1972] (1972)   
- "Resource-Related Research - Computers and Chemistry" Annual Report, 1973 (October 12, 1973)   
- [ACME Annual Report, 1973] (1973)   
- "Resource-Related Research - Computers and Chemistry" Annual Report, 1975 (May 7, 1975)   
- [SUMEX Annual Report, 1975] (May 1975)   
- [SUMEX Annual Report, 1976] (May 1976)   
- "Resource-Related Research - Computers and Chemistry" Annual Report, 1976-1977 (February 1977)   
- [SUMEX Annual Report, 1977] (June 1977)   
- "Resource-Related Research - Computers and Chemistry" Annual Report, 1977-1978 (February 1978)   
- [SUMEX Annual Report, 1978] (May 1978)   
- [SUMEX Annual Report, 1979] (May 1979)   
- Report of the Committee on the Graduate School (December 14, 1979) 
- Annual Report of the President, National Academy of Sciences, 22 April 1980 (April 22, 1980) 
- [SUMEX Annual Report, 1981] (June 1, 1981)   
- [SUMEX Annual Report, 1982] (June 1, 1982)   

Number of document segments

The Joshua Lederberg Papers

Title:
 The Next Move 
Number of Image Pages:
 5 (289,759 Bytes)
Date:
 1966
Creator:
 Lederberg, Joshua
Rights:
 Reproduced with permission of Joshua Lederberg.
Relation:
Lederberg Grouping: SAM Science and Man Series
Unique Identifier:
 BBGGFV
Accession Number:
 100
Document Type:
 Essays
Language:
 English
Format:
 application/pdf
 image/tif
Physical Condition:
 Good
Series: [Writings](#)
 SubSeries: [Science and Man](#)
 SubSubSeries: [Background & Editorial Files](#)
 Folder: [SAM Heading](#)
Metadata Last Modified Date:
 2000-12-18

Dublin Core of CHM Metadata

Dublin Core Metadata – an international standard:

title	author, designer, photographer, artist, architect, etc.
creator	
subject	publisher, manufacturer, etc.
description	
publisher	
contributor	text physical object image software sound moving image
date	
resource type	
format	
Identifier	
source	
language	
relation	
coverage	
rights	

Lederberg Artifact types

- [Abstracts](#)
 - [Agendas](#)
 - [Announcements](#) m;
 - [Application forms](#)
 - [Articles](#) m
 - [Autobiographies](#) m
 - [Bibliographies](#) m
 - [Biographies](#) m
 - [Brochures](#) m
 - [Certificates](#) m
 - [Correspondence](#) m
 - [Diaries](#) m
 - [Drafts \(documents\)](#)
 - [Drawings](#) m
 - [Electronic images](#) m
 - [Essays](#) m
 - [Eulogies](#)
 - [Excerpts](#)
 - [Grant proposals](#)
 - [Interviews](#) m
 - [Invitations](#)
 - [Laboratory notebooks](#) m
 - [Laboratory notes](#)
 - [Lecture notes](#)
 - [Lectures](#) m
 - [Legal documents](#) m
 - [Legislative records](#)
 - [Lists](#)
 - [Manifestoes](#)
 - [Memoirs](#) m
 - [Minutes](#)
 - [Monographs](#) m
 - [Narratives](#)
 - [Newsletters](#)
 - [Newspaper columns](#) m
 - [Notebooks](#) m
 - [Notes](#)
 - [Obituaries](#)
 - [Official reports](#)
 - [Oral histories](#) m
 - [Petitions](#)
 - [Photographic prints](#) m
 - [Press releases](#) m
 - [Procedures](#)
 - [Proceedings](#) m
 - [Programs](#) m
 - [Proposals](#) m
 - [Questionnaires](#)
 - [Reminiscences](#)
 - [Reports](#) m
 - [Resolutions](#)
 - [Resumes](#)
 - [Reviews](#) m
 - [School records](#)
 - [Speeches](#) m
 - [Summaries](#)
 - [Tables \(documents\)](#)
 - [Technical reports](#) m
 - [Transcripts](#) m
 - [Typescripts](#)
 - [Video recordings](#) m
- m= Computer Museum

Newell, Lederberg, MLB Content types for personal archives

Allen Newell: 65 types

Joshua Lederberg: 62 types

In Common: 24

MLB: 185 content types x 27 dissemination formats

MLB: 85 in common with AN or JL

MLB: 15 unique additions

MLB: 15 missing that either AN or JL have:

- too specific (Federal Register), encyclopedia,
- unclear e.g. lab notes
- to be added e.g. memoirs

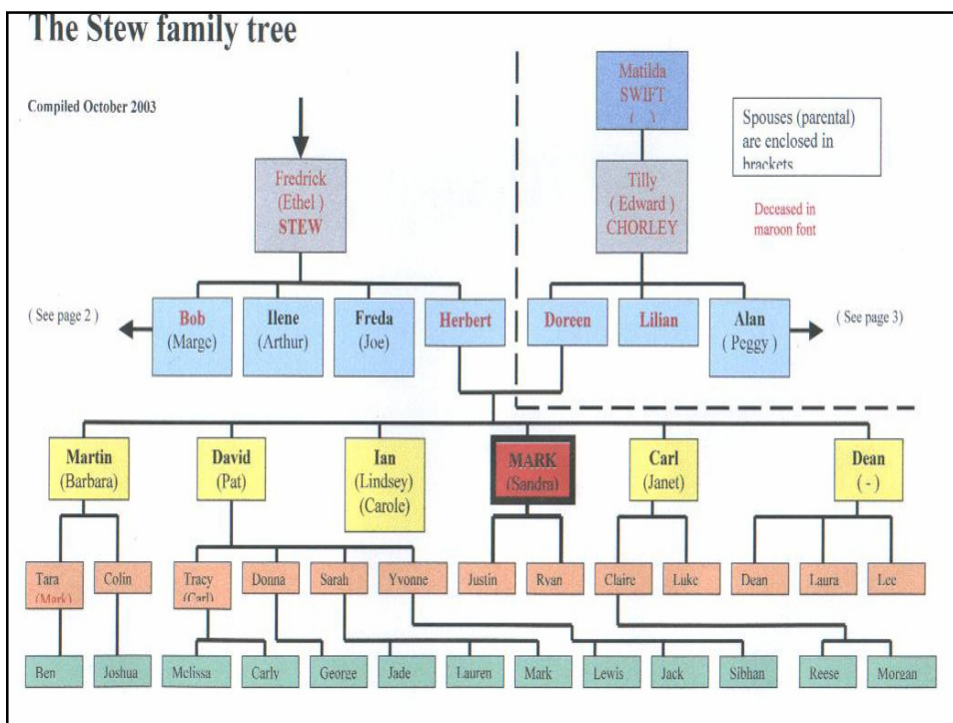
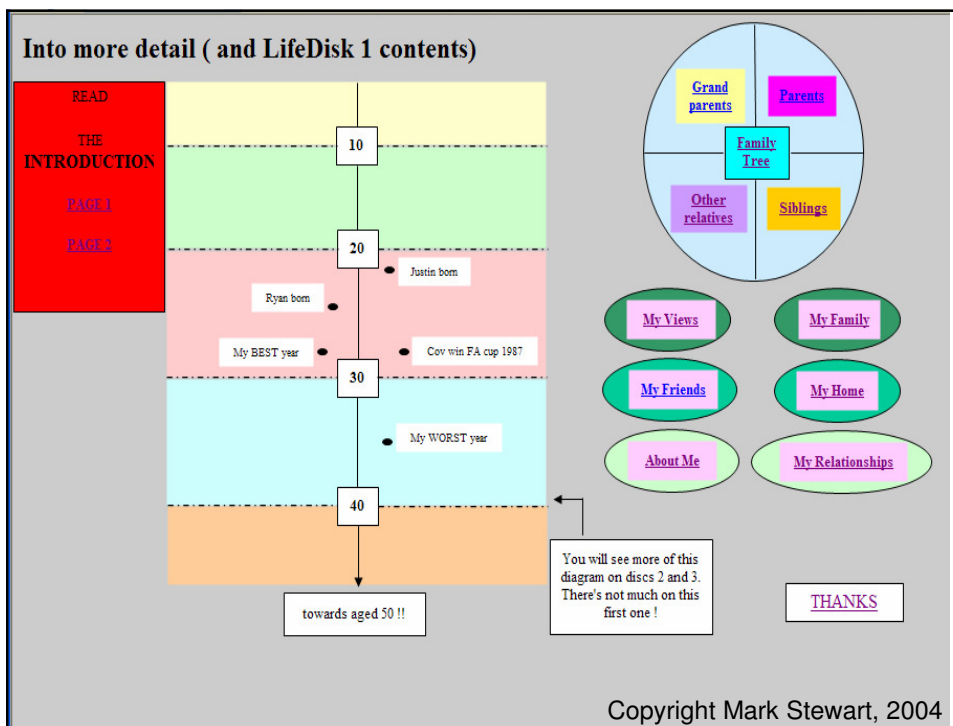
Lederberg artifact types

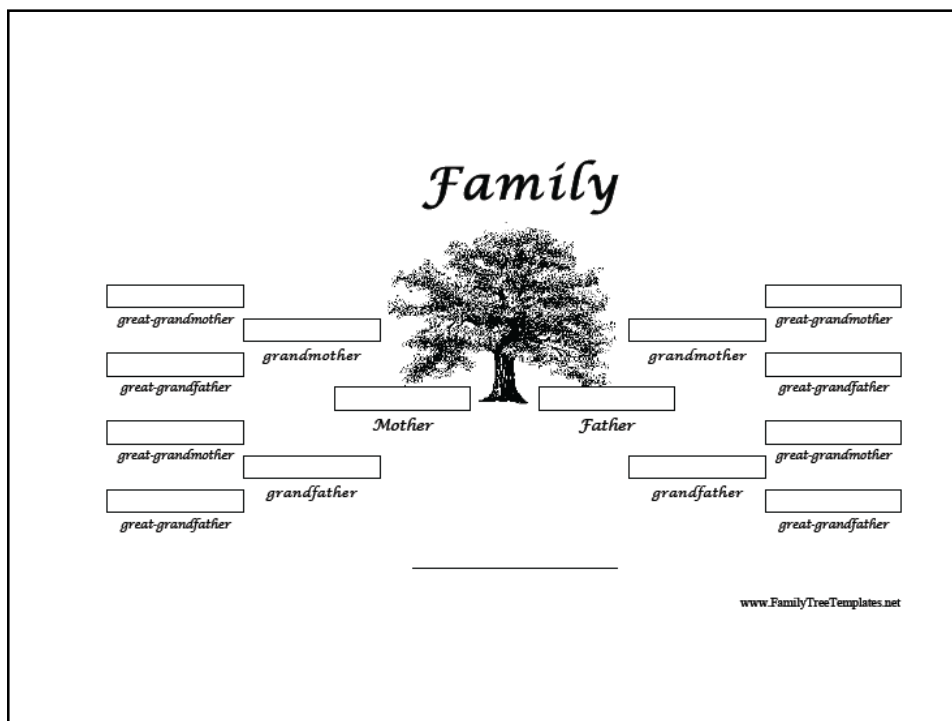
Which items will NOT be digital?

- | | | |
|---------------------------------------|--|-------------------------------------|
| • Abstracts | • Laboratory notebooks m | Press releases m |
| • Agendas | • Laboratory notes | Procedures |
| • Announcements m; | • Lecture notes | Proceedings m |
| • Application forms | • Lectures m | Programs m |
| • Articles m | • Legal documents m | Proposals m |
| • Autobiographies m | • Legislative records | Questionnaires |
| • Bibliographies m | • Lists | Reminiscences |
| • Biographies m | • Manifestoes | Reports m |
| • Brochures m | • Memoirs m | Resolutions |
| • Certificates m | • Minutes | Resumes |
| • Correspondence m | • Monographs m | Reviews m |
| • Diaries m | • Narratives | School records |
| • Drafts (documents) | • Newsletters | Speeches m |
| • Drawings m | • Newspaper columns m | Summaries |
| • Electronic images m | • Notebooks m | Tables (documents) |
| • Essays m | • Notes | Technical reports m |
| • Eulogies | • Obituaries | Transcripts m |
| • Excerpts | • Official reports | Typescripts |
| • Grant proposals | • Oral histories m | Video recordings m |
| • Interviews m | • Petitions | |
| • Invitations | • Photographic prints m | m= Computer Museum |

Mark Stewart

- I envy Mark's beautifully done, easy to navigate archive of his life. His family tree with hyperlinks to various relatives and lifelines provide a wonderful model for us all. I think I know Mark. In fact after receiving one of his DVD update shipments, I urged him to stop smoking.
- Mark Stewart is a software engineer from Coventry in the West Midlands of Great Britain. After reading about MyLifeBits in 2003, Mark designed and built what he calls his MyLifeDisk. It includes: over 4,000 files, 1,260 web pages, 1,082 pages of text, 2,100 photographs, 12 spreadsheets, 46 soundtracks and 72 video--seven gigabytes of autobiographical material.
- "Discover where he when to school and what it looked like plus school memorabilia. Where I went to college and what it looked like. Where I grew up and what it looked like. Old cine film footage of me during my childhood years. What toys I used to play with as a child. Where I've worked and what jobs I've had. What holidays I've had. Former girlfriends I've had. What hobbies and interests I've had or still have. What makes me laugh. Ambitions. What music I like. What films I like and have seen. What jewelry I wear and aftershaves I use. A family tree and view of my relatives. What friends I've had and still have—and our favorite pubs. In fact, almost anything about my life I've been able to record and find so far."
- Mark's life is organized in a set of hierarchical folders that sat the top level are
 - Family (folders for parents, siblings, ... grandchildren, and family tree for navigation), Homes,
 - Yin (education, interests, work),
 - Yang (humour), and
 - Self (folders for Friends, Relationships, Facts, Character traits, Interests, and
 - Memorabilia (with sub areas of Events and Social Life))






How far do we trust our institutions to save lives?

- Re a comment on NPR in late January <http://www.npr.org/templates/story/story.php?storyId=99372779> about people saving recordings of early pre-bluegrass American folk music:
- "He considered giving his collection to the Library of Congress, ... Alden says he worried that they'd be hard for musicians ... to access, and that they'd gather dust lying ..., what librarian ... would let someone into the stacks with a banjo or a fiddle ...?"
- ... they're burning CDs and shipping them all over, which is the "lots of copies keeps stuff safe" philosophy (www.lockss.com). They haven't taken the next step and put them online, and anyway don't have a virtual place to put them that has a good chance of surviving and caring for them in perpetuity.

What about 21st people?

- Born and living digital
 - Articles, books, two way communication (email)
- These individuals live in many cyberspaces or universes... web pages, lifelogs, blogs, etc.
 - Blogs, Facebook, MySpace, Twitter, ... Omega
 - (Who owns it? Can you get it? Do you want it?)
 - They might write with  that have no ink
- Highly personal Finances, health, specific email or letters, diaries

A View of Preserving Digital Lives

- Preserving the analog life of a 20th century person: 10-100 GB. 2 Mpgs, 100Kimages, 100-1,000 hrs. video
 - Won't analog people need to be converted to digital, ... if not they're really gone and forgotten history?
 - Gresham's Law: digital lives drive out analog lives
- How will a 21st century, digital person be preserved?
 - Which "lives" of a person e.g. personal, professional?
 - Depth of each life?
 - Size. Who's in the BL's digital lifeboat?
- Preserving Everybody?
 - Role of public institutions vs. the cloud for "all of us"

The end.

About Digital Lives at the British Library

Introduction

The recording, retention and storage, and transmission of cultural memory are changing.

As people increasingly create and share information via digital media, personal and cultural memory is being transformed.

For centuries, individuals have used physical artifacts as personal memory devices and reference aids. Over time these have ranged from personal journals and correspondence, to photographs and photographic albums, to whole personal libraries of manuscripts, sound and video recordings, books, serials, clippings and off-prints. These personal collections and archives are often of immense importance to individuals, their descendants, and to research in a broad range of Arts and Humanities subjects including literary criticism, history, and history of science.

Personal collections of scholarly interest to the Arts and Humanities tend to focus either on:

Particularly influential and eminent individuals for example the papers and correspondence of famous scientists, or

Literary authors or groups of individuals collated by organisations and projects for example the diaries in the Mass Observation Archive at Sussex University or the 6,000 oral history interviews in the Millennium Memory Bank now held in the British Library.

These personal collections support histories of cultural practice by documenting creative processes, writing, reading, communication, social networks, and the production and dissemination of knowledge. They provide scholars with more nuanced contexts for understanding wider scientific and cultural developments.

As we move from cultural memory based on physical artifacts, to a hybrid digital and physical environment, and then increasingly shift towards new forms of digital memory, many fundamental new issues arise for research institutions such as the British Library that will be the custodians of and provide research access to digital archives and personal collections created by individuals in the 21st century.

The Digital Lives Project

Digital Lives is a major research project focusing on personal digital collections and their relationship with research repositories. It brings together expert curators and practitioners in digital preservation, digital manuscripts, literary collections, web-archiving, history of science, and oral history from within the British Library (one of the world's leading research libraries) with researchers in the School of Library, Archive and Information Studies at University College London, and The Centre for Information Technology and Law at the University of Bristol.

We expect the findings on personal digital collections from the Digital Lives project will be applicable and of considerable interest to a wide-range of individuals and research repositories. These materials form a substantial part of the collections of national and regional libraries and archives which are widely used by academic staff and students. We therefore expect outcomes from our research to be of significant interest within the Arts and Humanities and the libraries, archives, and information sector. It will also be of potential interest to researchers exploring applications of digital memory in other areas such as health and ageing populations and for individuals who wish to manage their own personal digital collections for family history or other purposes.

The project is considering not only how collections currently being deposited are changing but also the fate of the research collections of the future being created now and implications for collection development and practice.

We are addressing these aims by focusing on the following objectives:

- Establishing how modern personal digital collections are being created, managed, and made accessible (either now or in the future)
- Exploring the needs and views of potential scholarly users of future personal digital collections such as biographers and historians
- Establishing how curators should approach selection, description, preservation and access for these personal digital collections
- Identifying and discussing with potential donors and curators the implications and methods of transfer of personal digital collections from individuals to long-term repositories
- Researching and assessing potentially transferable tools such as computer forensics or synchronisation which could assist with management, description, and preservation of personal digital collections or facilitate access to them
- Establishing the impacts of legislation, confidentiality and professional ethics on personal digital collections and the implications for their dissemination, or acquisition by repositories
- Identifying how we should address "hybrid" personal collections of digital and traditional media
- Researching new services acting as intermediaries for managing and publishing personal digital collections and assessing their potential effectiveness for preservation
- Sharing our experience and outcomes from this research with others and help build a community of repositories and researchers working in this area. Our overall aim is to make a substantial impact in this field of research and providing the future foundation for larger scale research projects.
- Further information on our methodology and work packages can be found in our Research Proposal (Case for Support) on this website.

Project Board

The project is directed by a project board of senior staff consisting of Prof. David Nicholas and Ian Rowlands (University College London), Neil Beagrie, Katrina Dean and John Tuck (British Library).

The principal investigator Neil Beagrie is responsible for overall management of the project and first point of contact for all exchanges with the AHRC. The co-investigators Katrina Dean and Ian Rowlands co-ordinate and supervise respectively the curatorial work and investigative survey.

External Project Advisory Board

An external board provides input to the development of the project and acts as external advocates of its work. Members are:

Sheila Anderson (Arts and Humanities Data Service)
 Prof. Wendy Hall (Southampton University)
 Christian Lindholm (Fjord)
 Cliff Lynch (Coalition for Networked Information)
 David Thomas (The National Archives)
 Susan Thomas (Oxford University Library)
 Prof. Jonathan Zittrain (Oxford University/Harvard University)

INTRODUCTION TO THE CONFERENCE

Helping to enable worldwide and interdisciplinary collaboration in securing personal digital archives for all people whose lives are increasingly digital.

Research Conferences 2.0

The first Digital Lives research conference will be a participatory and collaborative one, gringing together researchers, professionals, creators, enthusiasts and the digital public. The conference will be engaging colleagues beyond the project, and individuals in the broader community, in the research process itself.

Participants will be invited to contribute to the project. We shall be seeking feedback, suggestions, opinions and experiences from all those in attendance.

Perspective

The origin of personal computers in the mid 1970s and the arrival of the graphical user interface in the 1980s were instrumental in the evolution of personal digital archives; but it was the increasing integration of these computers within networks, most especially the internet, that yielded the most radical and still emerging changes in personal archives. The arrival of email and the world wide web transformed the way people communicate, and the widespread ability – gained in the last 5 years or so – to readily store and transfer files sufficiently large for high resolution images and sounds and moving images means that the nature of personal archiving has changed for ever.

No longer can the family documents, photographs and other memorabilia merely be boxed and stacked in the attic. At the same time, there are new opportunities for personal and family memory and for new uses of personal archives.

Themes

In future, Digital Lives conferences will likely want to focus on specific themes and particular aspects of this new research field. For this first conference the aim is to take the broadest perspective in in order to map out the issues, the challenges and, above all, the opportunities. A central question of this conference is how can repositories – national, regional and local - help academics and all people whose lives are becoming increasingly digital, to secure, preserve and organise their personal digital archives for themselves and for their families. In some cases this might ultimately be in readiness for future transfer to a repository where appropriate; in others it might be for longterm holding by the family.

Initially these archives will be for individuals' own use. It is anticipated, however, that people who feel in control of their personal archives - robustly secure and well organised may choose to make some of the contents of their personal digital archives available for research and other uses, sometimes in anonymous form.

The conference seeks to help enable – through the use of personal archives – the continuing contemplation and better understanding of personal lives by current and future generations, in accordance with the varying wishes of individuals concerned. Digital Lives Research Conference **DRAFT TEXT v1.2 12.i.2009**

Three Days in February

Day 1 Synthesis for Digital Lifelines: Practicalities, Professionalities and Potentialities

Day 2 Personal Information Lifecycles: Perspectives from Creator, Curator, Consumer

Day 3 Living Online and Personal Digital Archives in the Wild

PROGRAMME OF EVENTS (AN EVOLVING TIMELINE AND AGENDA)

Day 1: Digital Lifelines: Practicalities, Professionalities and Potentialities

Registration opens at 9:30 am with the Welcoming Address beginning at 10:00 am

- **Session 1: Demonstrations & Practical Experiences**
- **Session 2: Aspects of Digital Curation and Digital Preservation**
- **Session 3: On the Monetary Value of Personal Digital Objects**
- **Session 4: Open Source with a Keynote Lecture by Chris DiBona of Google**
- **Session 5: Professional Matters Arising and Resolutions**
- **Session 6: Options for the Future**

International Speakers include:

John Blythe (University of North Carolina)

Erika Farr (Emory University)

Juan-José Boté (University of Barcelona)

Cal Lee (University of North Carolina)

Cathy Marshall (Microsoft Research)

Naomi Nelson (Emory University)

Michael G. Olson (Stanford University Libraries)

Ludmila Pollock (Cold Spring Harbor Laboratory)

Gabriella Redwine (Harry Ransom Center, University of Texas at Austin)

Pelle Snickars (Swedish National Library)

William Snow (Stanford University Libraries)

We hope to telecommunicate with Australian colleagues during the conference, helping to reduce our carbon footprint.

Day 2: Personal Information Lifecycles

This is a draft agenda and the order of presentations is subject to change

Registration opens at 8:30 am with the first address beginning at 9:15 am

Opening Keynote Address

Dame Lynne Brindley, Chief Executive of the British Library

Session 1: Digital Lives**Brief Introduction to Digital Lives**

Jeremy Leighton John, The British Library

The First Digital Lives Conference Keynote Lecture

Gordon Bell, Microsoft Research

Session 2: Creators' Experiences, Anticipations and Thoughts**Digital Lives Presentation**

Jamie Andrews, The British Library

Writers in ConversationSession 3: Personal Information Management and Usability**Digital Lives Presentation**

Ian Rowlands and Peter Williams, SLAIS, University College London

Usability Keynote Lecture

Victoria Bellotti, Xerox PARC, California

Session 4: Forensics, Authenticity, Security and Digital Capture**Digital Lives Presentation**

Jeremy Leighton John, The British Library

Forensics & Digital Capture Keynote Lecture

Simson Garfinkel, Naval Postgraduate School, California

Session 5: Historical Research and Private Lives**Digital Lives Presentation**

Robert Perks, The British Library

Historical Keynote Lecture

Orlando Figes, Birkbeck College London

Session 6: Scientific Research with People**Developmental Psychology Keynote Lecture**

Charles Fernyhough, University of Durham

Session 7: Towards Digital Biography**Literary Biography Keynote Lecture**

Andrew Lycett, London

Scientific Biography Keynote Lecture

Georgina Ferry, Oxford

Session 8: Legal and Ethical Issues**Digital Lives Presentation**

Andrew Charlesworth, CITL, University of Bristol

Privacy Keynote Lecture

Kieron O'Hara, University of Southampton

Closing Remarks

- **Day 3: Digital Archives in the Wild**
- The following sessions are anticipated
- Session 1: Collaborative iLITERATURE and iART
- Session 2: iSCIENCE with personal information, and Online Notebooks
- Session 3: Digital Economics, Reputation and Digital Rights Management
- Session 4: Emerging Technologies including Ubiquitous Computing
- Session 5: Future Access and Visualisation
- Session 6: Virtual Worlds
- Session 7: Digital Conversazione

- Jon Crowcroft, Marconi Professor of Networked Systems, University of Cambridge
- Timo Hannay, Publishing Director of Nature.com
- Ian Hughes of IBM's Hursley laboratory
- George Oates, a founder member of the team that built Flickr
- Ren Reynolds of the Virtual Policy Network
- Dave Taylor of the Medical Media and Design Laboratory, Imperial College London
- Day 3 will involve the participation of visitors and inhabitants of Second Life

Content type (5)	Document type (180)
<ul style="list-style-type: none"> • Audio <ul style="list-style-type: none"> - Clip - Man made <ul style="list-style-type: none"> • Industrial • Sound effects - Natural <ul style="list-style-type: none"> • Bird • Water <ul style="list-style-type: none"> - Running - Waves - Song - Voice <ul style="list-style-type: none"> • Answering machine • Comment or annotation • Conversation • Meeting • Memo or dictation • Presentation • Speech • Telephone cal • Document type (180) • Graphics <ul style="list-style-type: none"> - Art <ul style="list-style-type: none"> • Album art • Cartoons • Drawings • Paintings - Diagrams - Graphs - Maps - Schematics • Photograph <ul style="list-style-type: none"> - People <ul style="list-style-type: none"> • Group • Portrait - Scenery <ul style="list-style-type: none"> • City • Domestic • Rural • Video <ul style="list-style-type: none"> - Commercial - Home 	<ul style="list-style-type: none"> • Advertisement (5) • Ballot • Brochure • Business card • Calendar • Certificate (14) achievement, birth...diploma... license...stock, warrantee • Commentary @ analysis (4) annotation, essay, paper rev., performance rev. • Computer source code • Correspondence (4) announcement, invitation, newsletter, recommendation • Description (1) Standard specification • Eulogy • Exam • Financial <ul style="list-style-type: none"> - Bill, statement, invoice (3) - Budget - Estimate - Financial report <ul style="list-style-type: none"> • Annual report • Expense report - Purchase order - Term sheet • Form (6) application, COA, questionnaire, renewal, survey, tax form & return • Grade card • Greeting card • List (15) bibliography, computer program output, database report, experimental results, grocery, inventory, laundry, menu, price, references, roster, service, statistics, web • Name tag • Newsgroup posting • Obituary • Plans (4) business, goals, proposal, schedule ((3) agenda, order of service, program) • Poetry • Presentation • Record of events (5) biography & auto biography, diary, news, progress report • Reference work dictionary, encyclopedia, web reference • Reservation • Resume • Rules and guidelines • Scrapbook • Script (dramatic) • Summary & syllabus • Ticket or stub • Transcript (computer chat, interview, speech)
<p>Content type, document type</p>	