The MyLifeBits Lifetime Store

Jim Gemmell, Roger Lueder and Gordon Bell
Microsoft Research
455 Market St., #1690
San Francisco, CA, 94105
{jgemmell,rlueder,gbell}@microsoft.com

ABSTRACT

Storage trends have brought us to the point where it is affordable to keep a complete digital record of one's life, and capture methods are multiplying. To experiment with a lifetime store, we are digitizing everything possible from Gordon Bell's life. The MyLifeBits system is designed to store and manage a lifetime's worth of data. MyLifeBits enables the capture of web pages, telephone, radio and television. This demonstration highlights the application of typed links and database features to make a lifetime store something that is truly useful.

Categories and Subject Descriptors

H.3.0 [**Information Storage and Retrieval**]: General
H.5.4 [**Information Interfaces and Presentation**]: Hypertext/ Hypermedia – *Architectures, Navigation, User issues*

General Terms

Management, Design, Human Factors.

1. INTRODUCTION

The MyLifeBits system [2,3] is designed to store and manage a lifetime's worth of *everything* – at least everything that can be digitized. MyLifeBits is inspired by Memex, a personal store envisioned by Vannevar Bush in 1945 for use by scientists [1]. Memex was envisioned to store documents, photos, and audio, with support for full-text search, voice/text annotations, and hyperlink creation. MyLifeBits supports capture, storage, management and retrieval of many media types, and logs as much usage data as possible.

We have used Gordon Bell's life for an experimental corpus. Everything possible from his past has been digitized, including: articles, books, cards, CDs, letters, memos, music, papers, photos, posters, paintings, presentations, home movies, videotaped lectures, and voice recordings. These are combined with media from his PC such as digital photos, email, and calendar events.

Methods to capture ever more of one's life are multiplying. MyLifeBits includes tools that allow Gordon to record every chat session, copy of every web page visited, and selectively record telephone, TV and radio. Figure 1 shows Gordon wearing DejaView cameras mounted on his hat and glasses. An attached device saves the previous 30 seconds of video when a button is pressed. There are software/hardware packages for recording meetings, and even devices to record body attributes such as heart rate and temperature.

The proliferation of capture methods has been accompanied by (and partially spawned by) the exponential increase in digital storage, especially hard drive capacity. Commodity hard drives will exceed one terabyte within a few years, leading to an era where lifetime recording is affordable, even for continuous video recording [2,3].

While the capture and storage capabilities for lifetime recording are exciting, improved tools are required to make a lifetime store useful. The MyLifeBits system is designed to facilitate browsing, searching, managing and enjoying one's lifetime store. In this demonstration we highlight how typed links and database features are keys to the usefulness of a lifetime store.

Figure 1 - Gordon Bell with cameras on hat and glasses

2. THE MYLIFEBITS SYSTEM

The MyLifeBits system supports capture, storage, retrieval, reporting, annotation, and story creation. Figure 1 illustrates the MyLifeBits system. The MyLifeBits data model includes various entities such as photos and documents, and typed links between entities. For example, a link from a contact to a photo may be of type "person in photo."

ETP'03, November 7, 2003, Berkeley, California, USA.

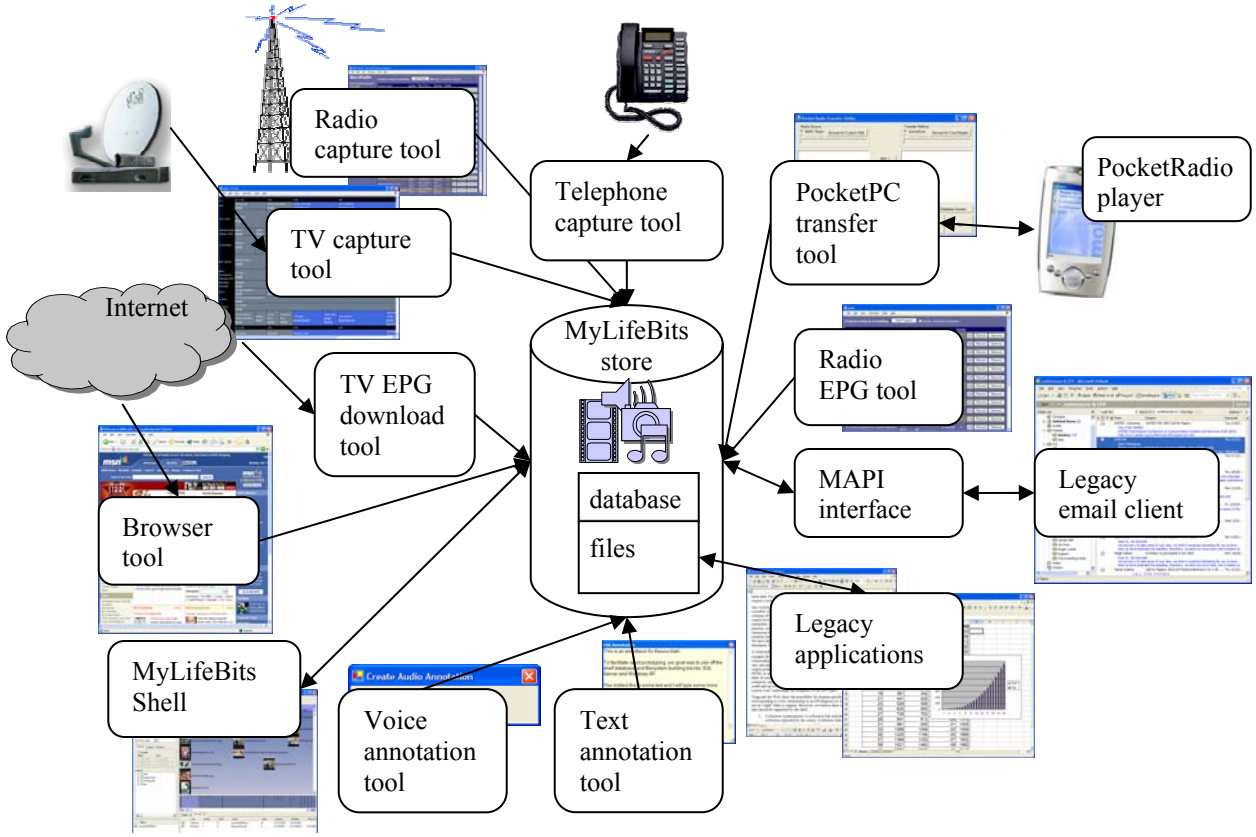


Figure 2 – The MyLifeBits system

We have discussed elsewhere the importance of annotations [2]. Annotation creation must be easy, in whatever mode strikes the user's fancy, and available at the moment the fancy strikes. In this demonstration, we perform voice and text annotations in a web browser, and in the MyLifeBits shell. Any number of selected items may be annotated in a single operation.

In MyLifeBits, the traditional folder (directory) tree is replaced by collections that form a directed acyclic graph (DAG). Any object (including a collection) may be filed in any number of parent collections (as long as cycles are not formed, violating the DAG constraint). Such filing should be made as convenient as annotation. For example, a button on the web browser toolbar files a web page in a "reading stack" collection for future reading.

While moving beyond the traditional folder hierarchy to a collection DAG is powerful and important, having links causes one to re-think many of the uses of collections. For instance, a collection for a date range is much better represented by a saved query for objects matching the date range. A collection to hold objects related to an event is better represented by links from the object to a calendar event object. Similarly, a collection of objects related to a person is best represented by links from the objects to the person object. Support for typed links can prevent misfiling. For example, one could accidentally file a photo in a collection called "phone calls with Sam", but it is impossible to create a "call from" link between a photo and the "Sam" person entity, because the "call from" link is typed to be from a person to a phone call only.

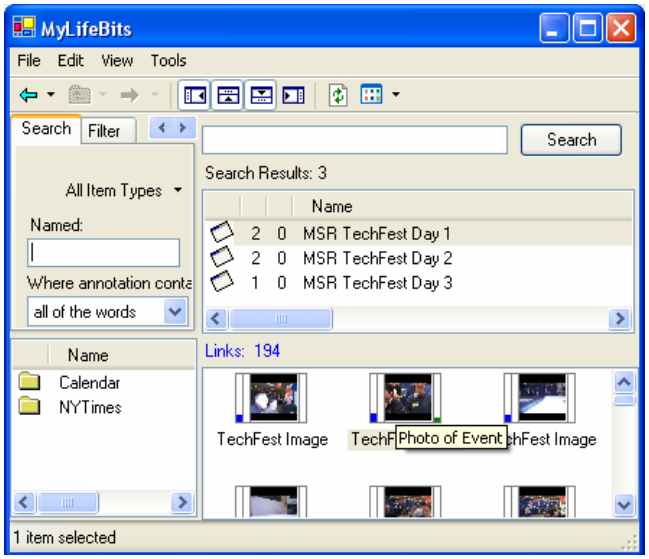


Figure 3 – Links pane (bottom right) shows instances linked to the selected item from the search results (top right). In thumbnail view, the link type is indicated by a popup on mouse hover. The bottom left pane shows collections the selected item belongs to.

The MyLifeBits UI shows all instances linked to the currently selected instance (Figure 3). In detail view, the link type is a column. In thumbnail and timeline view, the

link type is displayed as a pop-up when the mouse hovers over the item. When the selected instance is contained in other instances, they are shown in a special parent window pane.

Using a database provides fast search. User searches can be saved so that they can be repeated with a single click.

Additionally, the system, or third parties, can provide complex SQL queries beyond what a user could specify in any feasible form-based UI. For example, MyLifeBits includes a "commonly used files" query that considers the number of times a file has been used, n, with the time since last use, t, and the time span of its use, T (i.e. the time from its first use until the time of its last use). The sort order is based on the function

$$f(n,t,T) = (t/\alpha T)^{\beta} + 1/\gamma n$$

where α, β, and γ are constants. By appropriately setting the constants, one can capture the notion that a document that has been opened over the course of a year, and then not accessed for a week, is likely to be used again. However, another document that has only been opened over the course of two days, and then not accessed for a week, is not as likely to be used again.

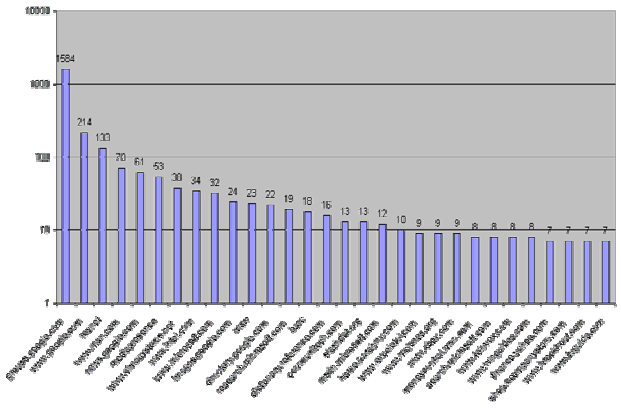


Figure 4 – MyLifeBits reports of most visited web sites

With a database, one can also create reports to understand what is in your lifetime store, and how it is used. Figure 4 shows a MyLifeBits report of the most commonly visited web sites. Figure 5 shows a report of email containing the text "funding" plotted versus time – something a user may want to see to get an idea of when busy times related to funding are happening. Seven standard reports are included in the current version of MyLifeBits, and, of course, any number could be added.

The indices in the database can be exploited to support pivoting, especially pivoting by time. That is, given some object, you can ask the system to pivot by time and show everything with a timestamp close to that of the object. The following example illustrates how hyperlinks and pivoting can be used to find objects that could not be found otherwise. Suppose that Gordon has a phone call with his realtor to discuss pricing his home, and his realtor tells him the URL to open on the World Wide Web to view a comparable property. Months later, Gordon remembers this property and wants to look at the page again. He knows it has been saved by the MyLifeBits browser tool. However, he cannot remember any text from the page to search for, nor does he remember when the call happened very accurately, just that it was in the fall. What Gordon can do is look up his realtor from his contacts. Selecting the contact for his realtor displays a window with all items linked to the realtor. One of these items is a phone call recording with the link "caller" to the realtor's contact (this link was created at the time of the call based on the caller-id). Gordon could listen to the recorded call to hear the URL, but that would be tedious. Instead, Gordon right-clicks on the call and selects "pivot/time-overlap". This lists everything in the database with a time falling inside the span of the phone call. As the web page was visited soon during the call, Gordon can easily scroll through the thumbnails of the web pages to find the desired page (which is probably no longer on the Internet – but Gordon has a copy in MyLifeBits).

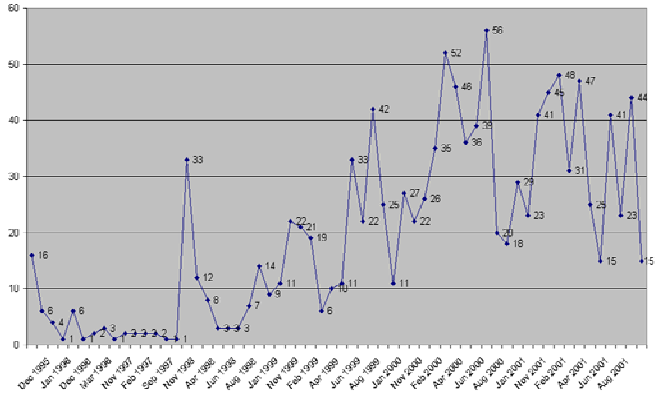


Figure 5 – MyLifeBits report: Mail received containing the text "funding" plotted versus time.

Usage logging can be applied in many ways. Figure 7 shows the event log for a movie. This log can be used to automatically detect highlights or sections that the user always skips. MyLifeBits supports a feature to play a movie with the same skipped sections as the previous play.

Use of a database also makes narrowing of search results very convenient and flexible. The MyLifeBits UI allows the user to filter results to only certain entities. For example, a search that returns photos, documents and videos, could be narrowed to only show the videos. We also perform clustering on some entity attributes for the purpose of filtering. For example, date attributes are clustered using a combination of largest-gap and k-means clustering (Figure 6). The user can click on a date cluster to narrow the search results. For text attributes like email subject or web page title, MyLifeBits shows the top seven

occurring strings as clusters and puts the rest in an "other" cluster (Figure 8).

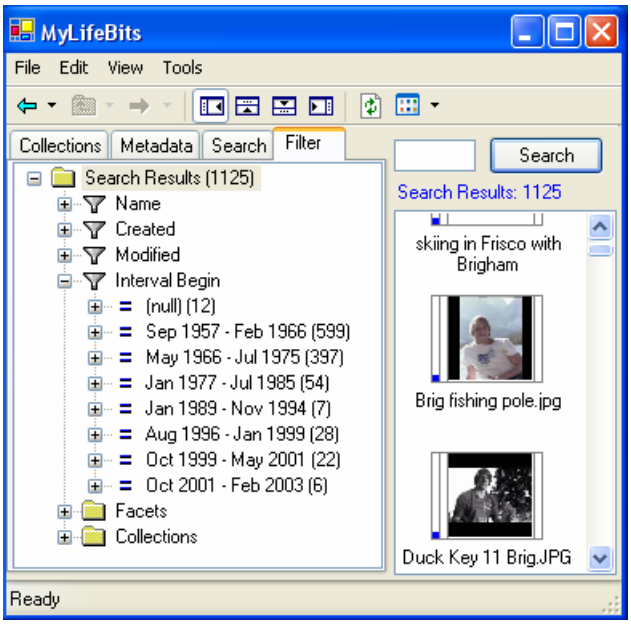
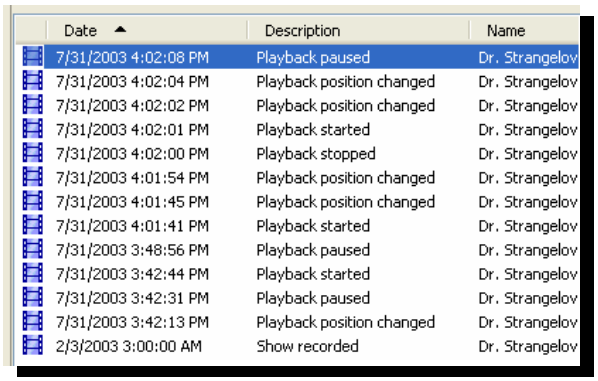


Figure 6 – Date clustering of search results using largest gap and k-means.



Date	Description	Name
7/31/2003 4:02:08 PM	Playback paused	Dr. Strangelov
7/31/2003 4:02:04 PM	Playback position changed	Dr. Strangelov
7/31/2003 4:02:02 PM	Playback position changed	Dr. Strangelov
7/31/2003 4:02:01 PM	Playback started	Dr. Strangelov
7/31/2003 4:02:00 PM	Playback stopped	Dr. Strangelov
7/31/2003 4:01:54 PM	Playback position changed	Dr. Strangelov
7/31/2003 4:01:45 PM	Playback position changed	Dr. Strangelov
7/31/2003 4:01:41 PM	Playback started	Dr. Strangelov
7/31/2003 3:48:56 PM	Playback paused	Dr. Strangelov
7/31/2003 3:42:44 PM	Playback started	Dr. Strangelov
7/31/2003 3:42:31 PM	Playback paused	Dr. Strangelov
7/31/2003 3:42:13 PM	Playback position changed	Dr. Strangelov
2/3/2003 3:00:00 AM	Show recorded	Dr. Strangelov

Figure 7 – Playback event log for a movie

These techniques for narrowing search can also be useful as a way to browse the full contents of MyLifeBits (equivalent to narrowing a search for everything).

We have implemented two programs that suggest links to the user. Our photo capture wizard suggests that photos taken when an event in the user's calendar occurred are in fact photos of the event, and should have the corresponding link created. Our telephone recording application suggests that a contact with a phone number matching the caller-ID phone number is participating in the call.

3. CONCLUSION

We have entered an era of virtually unlimited storage, in which modes of capture are proliferating. It is becoming convenient and affordable to record a lifetime of what one sees and hears, along with many new data source such as user logs and sensor data.

The challenge now is to make the stored material useful. MyLifeBits implements a number of features originally proposed by Vannevar Bush, such as links, annotations, and full text search. This demonstration highlights the importance of augmenting these ideas with typed links and database features. Visualizations, complex queries, and pivots enable the user to find or browse many items that would otherwise remain unused.

There are many aspects of MyLifeBits that remained to be studied, including: privacy (and scalable privacy specification), effective communication of my life to others, social issues and user interface issues.

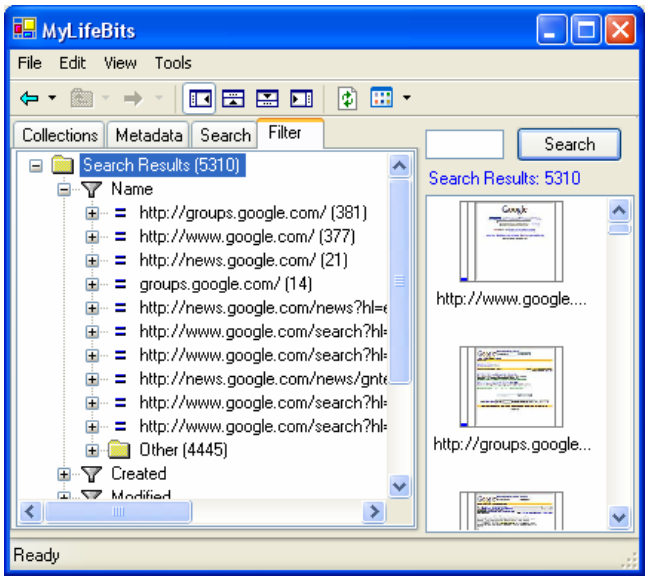


Figure 8 – Top occurring names in search result.

ACKNOWLEDGMENTS

MyLifeBits has benefited from the insightful comments of Jim Gray, Ted Nelson, and John Carlis. Josh Blumenstock and Evan Solomon wrote much of the code for the telephone, TV, and radio tools. Curtis Wong and Steven Drucker have helped us with UI issues.

REFERENCES

[1] Vannevar Bush. As We May Think, The Atlantic Monthly, 176(1), July 1945, 101-108.

[2] Jim Gemmell, Gordon Bell, Roger Lueder, Steven Drucker, and Curtis Wong, MyLifeBits: Fulfilling the Memex Vision, ACM Multimedia '02, December 1-6, 2002, Juan-les-Pins, France, pp. 235-238.

[3] Jim Gemmell, Roger Lueder, and Gordon Bell, Living With a Lifetime Store, ATR Workshop on Ubiquitous Experience Media, Sept. 9-10, 2003, Keihanna Science City, Kyoto, Japan, pp.69-76