

THE WORLD'S LARGEST UNDERSEA OBSERVATORY P. 38

SPECTRUM

IEEE

TOMORROW'S TECHNOLOGY TODAY

NOVEMBER 2005

www.spectrum.ieee.org

GORDON BELL

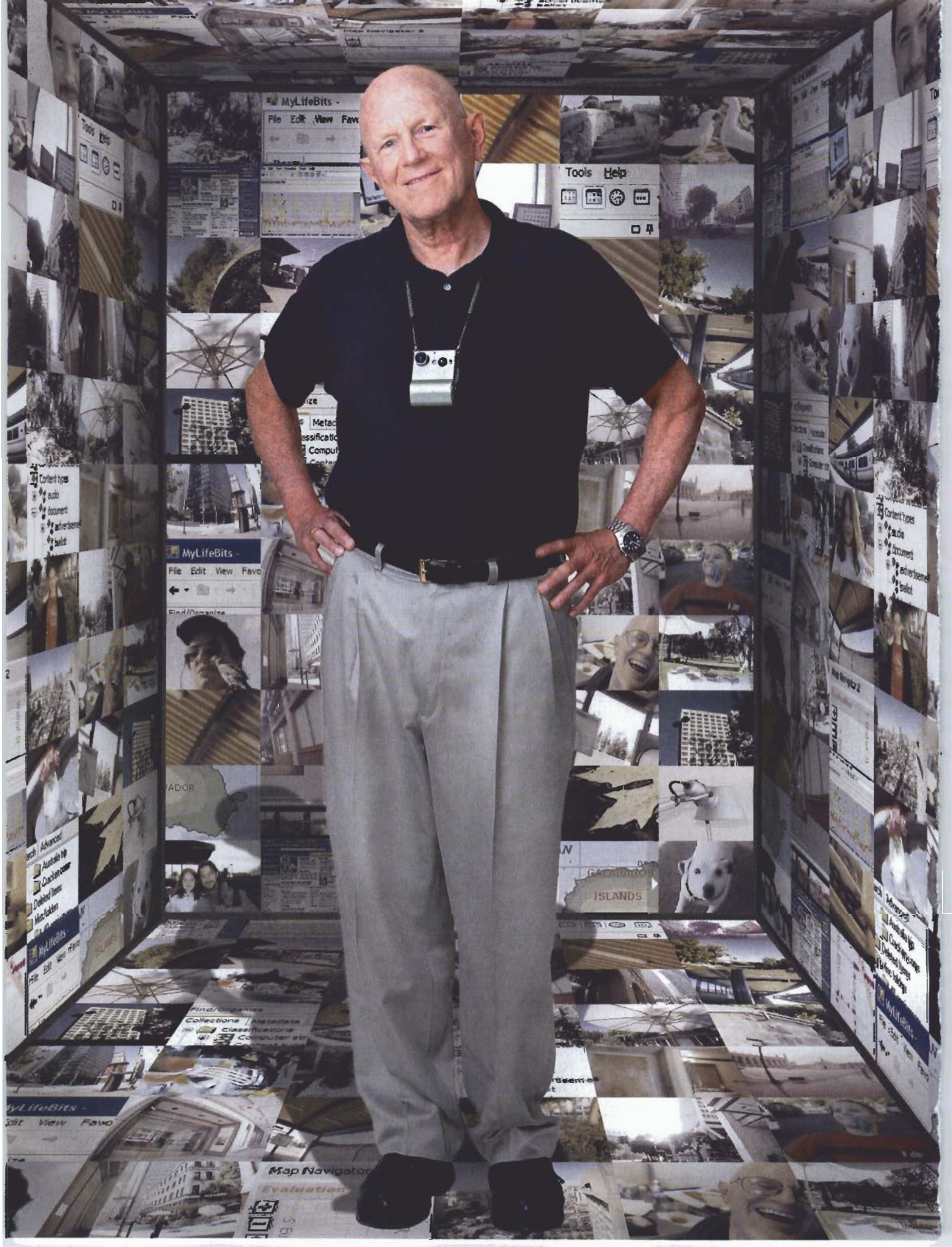
A LIFE, DIGITIZED

WITH NEW DATABASE SOFTWARE AND A WEARABLE CAMERA, COMPUTER PIONEER **GORDON BELL** IS HELPING MICROSOFT REINVENT THE PC AS **A PERSONAL MAINFRAME** FOR STORING ALL YOUR DAILY TRANSACTIONS

DISPLAY UNTIL: 3 December 2005
www.ieee.org US \$4.99 CAN \$8.50



IEEE



At age 71, computer pioneer Gordon Bell is as active, combative, and restless as ever. The legendary creator of Digital Equipment Corp.'s seminal VAX line of minicomputers in the mid-1970s is now a sort of researcher-at-large for Microsoft Corp., working in a tiny San Francisco center devoted to research on databases large and small.

Like many of us, Bell occasionally forgets things. But unlike most of us, Bell has both a tinkerer's heart and the wherewithal to launch a major software research project to make forgetfulness itself, like vinyl record collections and coffee-stained address books, yet another encumbrance remedied by the computer age.

His project, MyLifeBits, is the digital distillation of, almost literally, his every waking minute. It started out as an offhand experiment, but today its goal is nothing short of changing the way we use computers, and by extension, the way we live. At its heart, MyLifeBits is a big database on a personal computer, into which go the correspondence, keyboard-based chores, and

trans-Atlantic flight last year? Where is that Web site with a schematic and parts list for a radio-controlled clock? You saw it a few weeks ago. Which little hotel in Hong Kong did your brother-in-law recommend? What was that nice wine at the sales-meeting dinner last June? Countless details such as these, some pivotal, others trivial, could be as easy to call up as the names in that dog-eared address book.

What's in it for Microsoft? The Redmond, Wash.-based giant sees a way to straighten out our increasingly information-tangled lives—and an opportunity to reinvigorate the sluggish PC market. Take the MyLifeBits software, the SenseCam, and the low cost of memory of all kinds, which keeps plunging, with no bottom in sight. Mix in Microsoft's considerable marketing and research clout and you begin to envision the PC being reborn as the personal mainframe—a terabyte repository of all our life bits. And its indispensable data-collecting accessory, a commercialized version of the SenseCam—a camera-enhanced

TOTAL RECALL

A MICROSOFT RESEARCHER IS DETERMINED TO RECORD EVERYTHING ABOUT HIS LIFE.

EVERYTHING

BY STEVEN CHERRY

even the sights and sounds of everyday life. It automatically swallows up and indexes e-mails, keystrokes, recorded phone calls, images, video, and every Web page that graces its user's computer's screen.

For the past two years, MyLifeBits has been capturing real life with an unobtrusive miniature still camera that Bell [left] wears around his neck, pendantlike. Clever sensors detect light, heat, and position. Software tells the SenseCam, as the device is called, whether to snap a picture. At day's end, MyLifeBits grabs all those images and the sensor readings too. In short, just about everything that is digital, or that can be digitized, goes into the database—and is easily recalled weeks or years later. Researchers elsewhere have been capturing life bits for decades, but until Microsoft got in the game, no one had put so many varieties of them into a searchable and indexable database. The ability to sift so much information is impressive and, to many, intimidating.

If privacy is your overriding concern, MyLifeBits may not be for you. But if managing the details of a full and hectic life is a problem, the benefits of the system would be undeniable. Who was that venture capitalist you sat next to on that

cellphone, perhaps—will be made intelligent by an operating system designed in, yes, Redmond, Wash.

Bell's MyLifeBits project is now the epicenter of a far-reaching initiative at Microsoft that goes by the name "memory augmentation." At last count, three researchers were assigned to the software and three more to the camera. At least a dozen more spend some or all of their time extending the software's capabilities. The research is intertwined with a number of other projects and goals at Microsoft, making it impossible to say precisely how much R&D funding is involved. But make no mistake about its importance to the company—and to computer users everywhere. Consider just one part of memory augmentation, the area of computer science known as information retrieval—essentially the ability to look through a data collection and find the right item. At the field's top conference this year, Microsoft researchers accounted for more than 20 percent of the entire technical program of 71 papers.

Information retrieval is also where MyLifeBits research will first show up in Microsoft Windows Vista (aka Longhorn), the long-awaited major revision of the omnipresent Windows

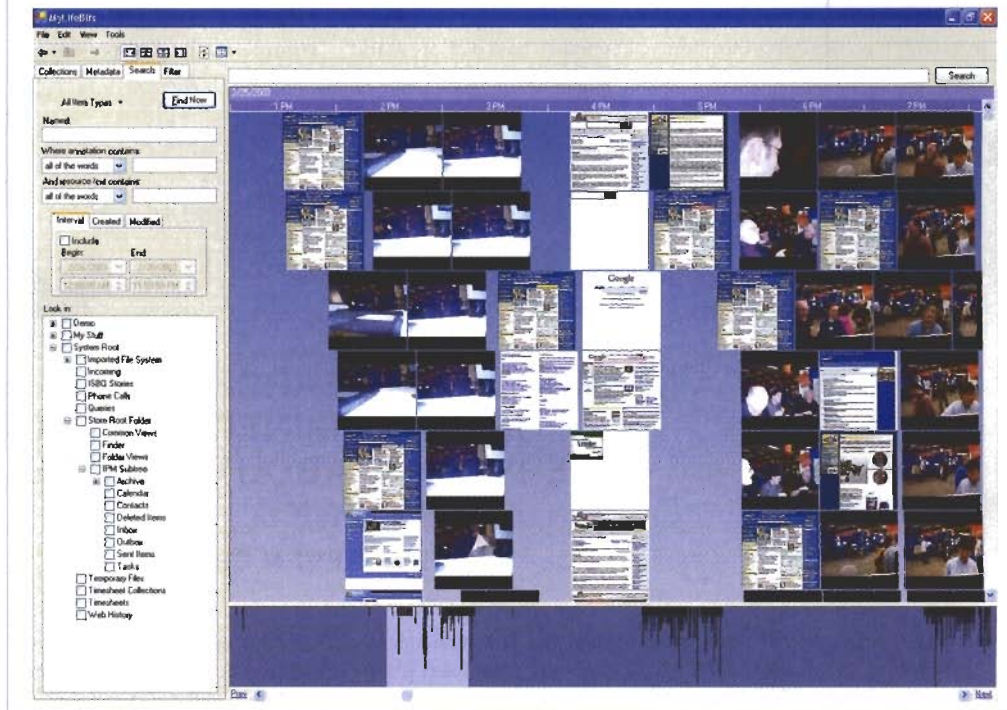
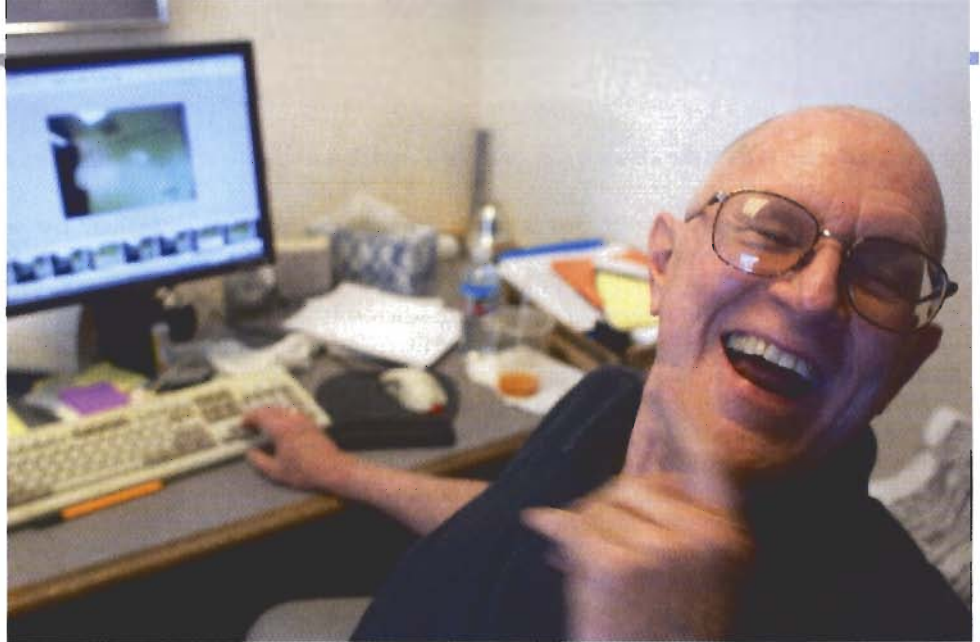
operating system. By all accounts, Vista will put search and retrieval front and center. The metaphor of files and folders—the central way we've been interacting with computers for almost 30 years—will be largely hidden from view [see "Interface Lift," in this issue]. Instead, the operating system will index all your documents, try to guess which ones are important to you at the moment, and let you intelligently browse and search through them if it guesses wrong.

This emphasis on search puts Microsoft on a collision course with its archrivals, Google and Apple. Each is racing to find new ways to put your most useful information—whatever that is right now—at your fingertips. Apple's latest operating system, Mac OS X Tiger, has a new way of searching a computer, called Spotlight, that quietly keeps an index of everything on the hard disk. Then, when you start a search, it quickly serves up everything that's relevant, whether it's an e-mail, office document, or PDF file. Meanwhile, equally ambitious software is in the works at Google, the company that has made its own name a synonym for searching. A new application, called Google Desktop, lets you "google" your hard disk in a way similar to a Spotlight search.

With Vista not slated for release until some time in 2006, Microsoft is already about a year behind Apple and Google. But the company is counting on its new search-centric operating system to leapfrog its competitors. If it succeeds, the loss of a year won't matter—the race for the best ways of collecting and retrieving information is clearly a marathon that will dominate the next decade of computer science and information technology.

Just as Microsoft is not alone in coming up with new search technologies, others, notably Steve Mann, are also creating gadgets that capture life bits. For years, Mann, who teaches at the University of Toronto's department of electrical and computer engineering, has spent week after week viewing, as well as recording, reality through a wearable video camera of his own design.

But all these efforts, in hardware and software—the Steve Manns on the one hand, and the Apples on the other—have focused on only half of the problem, either recording the minutes of our lives, or searching our desktops for useful minutiae already buried there.

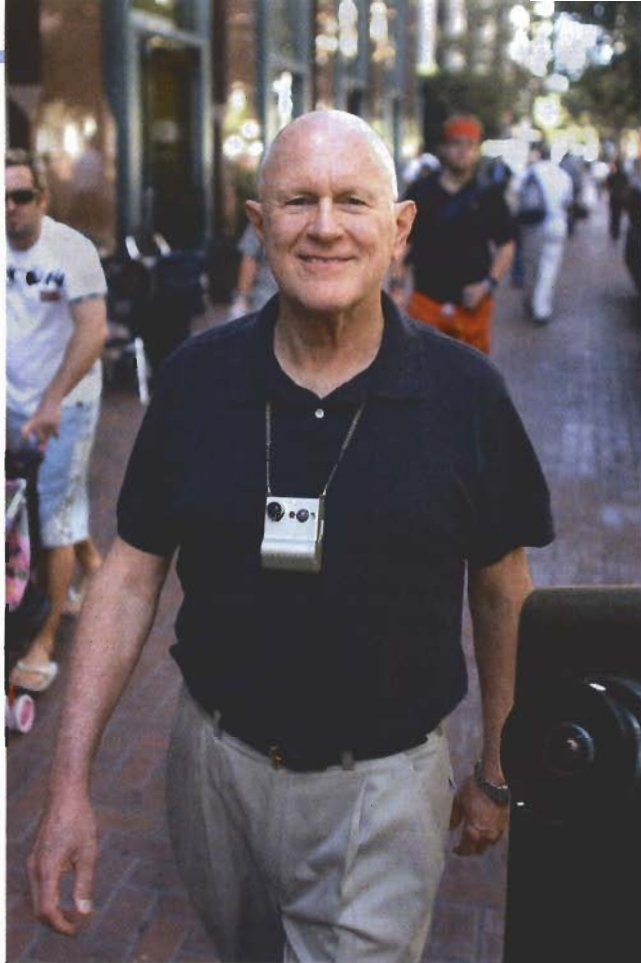


Only MyLifeBits has a binocular vision of creating digital doppelgängers for our entire lives, off-line as well as on.

INDEED, MYLIFEBITS is really the culmination of a 2700-year historical trend. Ever since the Greeks took the *Iliad* out of the singsong voices of traveling balladeers and set it down on clay tablets, the world's troves of recorded information have swelled relentlessly. Not surprisingly, humankind's strategy has been to rely less and less on actually remembering, and more and more on being able to find something when we need it. The Gospels, Gutenberg, and Google have all been steps along the way.

In 1945, Massachusetts Institute of Technology professor, Raytheon Co. cofounder, and presidential science advisor Vannevar Bush set out a blueprint for the electronic age, in an *Atlantic Monthly* article titled "As We May Think." Foremost among its astoundingly prescient ideas is a universal look-up machine Bush called the Memex, "a device in which an individual stores all his books,

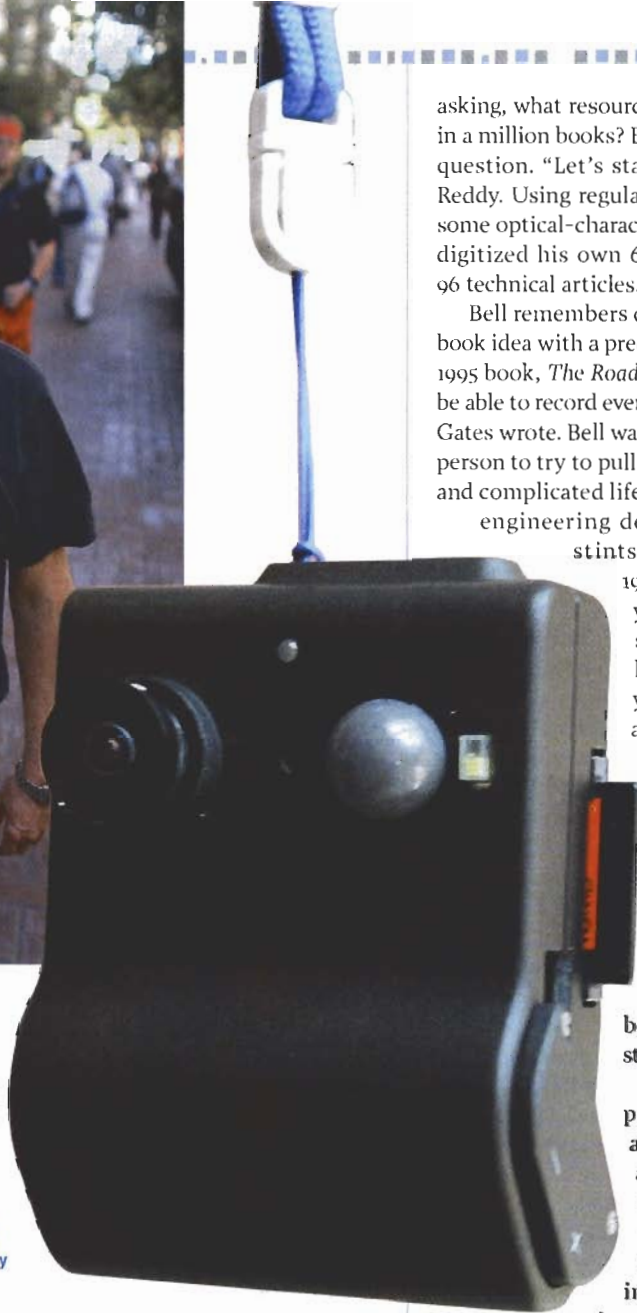
TOP: MARK RICHARDS; BOTTOM: MICROSOFT RESEARCH



BIT BY BIT: Microsoft researcher Gordon Bell records just about everything in his life in a database called MyLifeBits.

Sitting at his computer [top left], his every Web page viewed, document created, e-mail sent or received—indeed, every keystroke and mouse click—is saved. Away from his desk [above], Bell records daily life with a SenseCam [right], which takes pictures every minute or so and notes their time of day and GPS coordinates.

Back in the office, SenseCam photos and notations are downloaded to MyLifeBits [bottom left], which can call up bits of Bell's life in a variety of ways. Here, a search query is answered with thumbnails of photos and documents. A traditional display showing files and folders is relegated to only one part of the screen.



asking, what resources would it take to scan in a million books? Bell was delighted by the question. "Let's start with mine!" he told Reddy. Using regular desktop scanners and some optical-character-reading software, he digitized his own 6 books, 7 patents, and 96 technical articles.

Bell remembers connecting the million-book idea with a prediction from Bill Gates's 1995 book, *The Road Ahead*. "Someday we'll be able to record everything we see and hear," Gates wrote. Bell was in many ways the ideal person to try to pull off such a feat. His long and complicated life includes two electrical engineering degrees from MIT, two stints at Digital during its 1970s glory days, and six years teaching computer science at Carnegie Mellon. Then there were two years in the late 1980s as a director at the National Science Foundation, where he was responsible for the NSFnet, shortly before it was handed off to the private sector and became the Internet of today. He's also been an angel investor and a board member of more start-ups than he can recall. After scanning his published works, Bell added all his other files and papers as well, including ones from his time as head of engineering at Digital during the development of the VAX minicomputer.

Documents already in electronic form were funneled in as well, such as e-mail going back

records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is," he wrote, "an enlarged intimate supplement to his memory."

The 60-year-old manifesto is a nearly complete specification for MyLifeBits. The project's three major components are anticipated by Bush's descriptions of a head-mounted camera to capture information, annotations to help find data after the fact, and ways of organizing a collection, such as links from one document to another. Such a system has been something of a holy grail among information technologists for more than half a century. Microsoft may become the first company to pull it all together.

For Bell, the project began modestly enough seven years ago when a friend, Raj Reddy, asked about electronic copies of some of Bell's books and papers. Reddy, a distinguished computer science professor at Carnegie Mellon University, had in mind an idea that would become known as the Million Book Project. Reddy didn't intend to digitize a million books himself. The library project was a way of

to the late 1980s. PowerPoint presentations, digital photographs, drafts of articles, expense reports, phone bills—in it all went. His medical records, replete with gene sequences, would eventually get thrown in as well. Bell brags he's been "paperless" for three years now. What little paper is left in his life goes into a tray. "When the pile gets about two inches high, it gets scanned" and then discarded.

The project took on a new dimension in 2004, when a Microsoft engineer in the company's Cambridge, England, research lab developed a new wearable camera. It takes hundreds of pictures a day, automatically, without having to be held or aimed. Bell got his hands on one and started wearing it. Into the database those photos went, each with a date and time stamp. He also began wearing a GPS device to record where he was throughout his day.

The work took on a life of its own. Almost compulsively, Bell's coterie in San Francisco grabbed just about everything in Bell's life made of 1s and 0s. Today, every keystroke on his computer is captured as it is tapped. Office phone calls are not only logged but also

digitally recorded in their entirety. Every Web page is stored when it's viewed—not just the URL, but the entire image is jammed in there. [See photos, "Bit by Bit."] Data storage is dirt cheap—a gigabyte's worth, or about 10 000 Web pages, costs about US \$1.25.

Bell's life is mirrored by that database now. His colleagues in the San Francisco lab rummage through its more than 80 gigabytes each day in the course of their work. Other Microsoft researchers, in Redmond, have copies of the database as well. Among them is the team designing the graphics and other interface features that will be the public face of the Vista operating system. MyLifeBits might be the most complete and diverse personal data set on any PC anywhere. If Vista engineers can make sure Bell can find what he's looking for on his desktop, you'll be able to find anything you need on yours, so the interfaces they come up with are then tested back in San Francisco. It's no exaggeration to say that dozens of Microsoft researchers in four locations on two continents are spending thousands of man-hours turning the next PC desktop into a Memex.

"THE CAMERA HOUND OF THE FUTURE wears on his forehead a lump a little larger than a walnut," Bush told his magazine readers [see photo, "Ahead of His Time"]. Brownie cameras, Polaroid, and commercial color film, all developments of the late 1940s, were several years away. The device Bell wears today—originally on his forehead, nowadays around his neck—is bigger than a walnut. But it does far more than take pictures; it's designed to know when to take them. It has sensors for light level, x-y tilt, temperature, acceleration, and infrared.

"The SenseCam makes decisions," explains Jim Gemmell, a researcher who works closely with Bell. "If the light level suddenly changes, maybe I went from one room to another, so let's take a picture. Passive infrared says maybe it's a person; take a picture," Gemmell says. Passive infrared detects warmth—sophisticated burglar alarms work by making the same inference. Some of these levels, which Gemmell calls "trigger points," needed fine-tuning. "If the accelerometer says it's shaking like crazy, the camera waits a while. At first we got a lot of blurry pictures."

Gemmell, a quietly charismatic 40-year-old software engineer with an infectious smile, has been thinking about how to store multimedia files in databases for more than a decade. In fact, it was the topic of his 1995 doctoral thesis. These days, he spends his time figuring out new ways to record, and then sift through, life bits.

One of Gemmell's favorite SenseCam demos is a sequence of shots taken from inside a refrigerator. Because the camera detects the change in lighting, as soon as the door opens, it snaps a picture, so the "video" is a constant stream of people reaching into the refrigerator, putting stuff in or taking it out. "There's a little girl," he says. "She looks around, sees no one is looking, and steals someone's lunch. It's coming up." A few seconds later, he points at the screen, laughing. "There's she is." Sure enough, a brown-haired girl, about 6, looks furtively to the side, then grabs something.

The very first SenseCam video was shot last year by its inventor, Lyndsay Williams, who works in the Cambridge lab. A biomedical electronics engineer by training, Williams designed music

synthesizers in the 1970s and some of the earliest sound cards for personal computers in the 1980s. That original video [see screen shot, "Away-From-Home Movies"] is a simple one of traveling around town, but her gadget presents an odd point of view. We're used to films showing us someone walking about, but this is the opposite—photos taken from, not of, the person. The unseen Williams walks down the cobblestone streets of King's College. She enters a tea shop, then a church, later a train bound for London.

Jim Gemmell has shown the video many times to demonstrate what the SenseCam does. "It's a quick replay of a day," he says. "You can easily burn through a thousand photos. Besides trigger values, you can set a threshold 'for sure at least take a picture every minute or every 30 seconds' or whatever."

There are only 12 SenseCams in the world today. A second version is in the works, but Bell has been using the current model off and on for almost two years. There are so few of them, he's often had to lend his to other researchers. For a while, it was on loan to a University of Southern California graduate student. "The problem I've had," Bell says, "is everyone wants to try it."

At 85 by 65 by 35 millimeters and 175 grams (including its two AA batteries), the SenseCam is almost exactly the size and weight



AWAY-FROM-HOME MOVIES: A SenseCam will take a photo every minute or whenever its sensors detect a change in the environment, such as lighting or temperature. A chronological sequence drawn from the 1000 to 2000 photos taken in an average SenseCam day is at the bottom. At the top is a plot of the data points captured by the various sensors.

as two full-sized iPods stacked one on top of the other, and about the same weight as just one. If it were much heavier, it would be too uncomfortable to be worn by the aged and the infirm, who will surely be among the first beneficiaries of augmented memory. In fact, Microsoft is already investigating ways this group can be helped by MyLifeBits. In April, two neurophysiologists affiliated with Addenbrooke's Hospital, in Cambridge, England, were given six of the precious SenseCams. The first went to a "Mrs. S.," who suffers from limbic encephalitis, a rare degenerative disorder of the neurons.

The device has given Mrs. S. back the memories that her disease had taken away. "In the evening, her husband downloads its data to a PC," says Ken Wood, assistant director of the Cambridge

lab and manager of the SenseCam project. "We've built a simple viewing program, and they play the stream of pictures together. He talks her through the day."

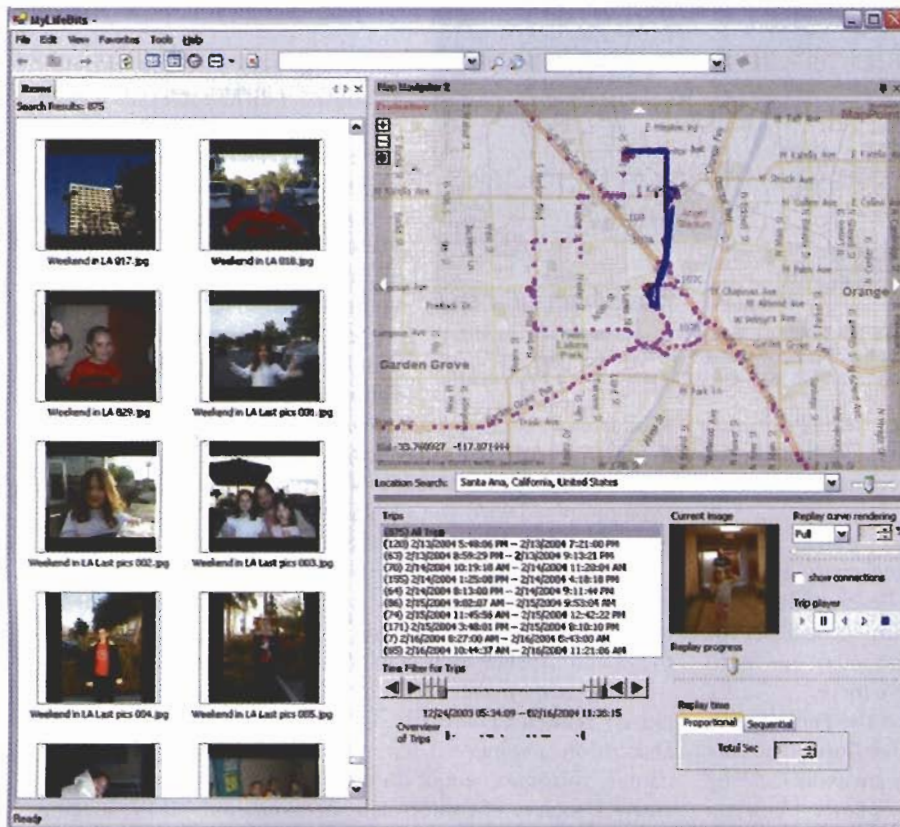
At 640 by 480 pixels (known as VGA resolution), the photos would look small and grainy and indistinct if printed. But strung together and viewed on a computer, they're a lifesaver for someone who often can't remember what she did a few hours earlier. Wood says that Mrs. S. used to be anxious throughout her day, because no matter how hard she tried, she would forget important things. "Now, she knows the SenseCam will remember for her," Wood says.

Video on a cellphone is at best 15 frames per second, and movie or television film is twice that. For the SenseCam, on the other hand, snapping one photo every 30 seconds maxes out its capac-

\$350,000 has been set aside for the winners, most of whom will be graduate students. Each will get the MyLifeBits database software, a few thousand dollars, and one of the scarce second-version devices. Microsoft has no immediate plans to commercialize the SenseCam. Only 100 will be made, by hand, in the Cambridge lab. The new model will be about the same size as the old one, but will do more, recording and storing audio data. And a GPS receiver will be built in, eliminating the need for the separate one that Bell had been walking around with.

The MyLifeBits team doesn't know to what new uses researchers will put the SenseCam, but they do know how useful it can be to have access to the minutiae of everyday life. Who was that venture capitalist you sat next to on that trans-Atlantic flight last year? Armed with the date and time of the flight (from your appointment calendar or even your expense reports), you can listen to the audio recordings made by the SenseCam. Sure enough, ten minutes into the flight, you introduced yourselves. What was that nice wine you had at the sales-meeting dinner last June? Probably at least one SenseCam photo taken at the time included the label.

Let's say you take a long drive up to the wine country [see screen shot, "Lost in Space"]. At the first winery, you realize you've lost your reading glasses. It was probably when you stopped for coffee. But where was that? Your SenseCam photos should be able to tell you, as they're really database records with a photo in one field, a time stamp in another, and a set of GPS coordinates in a third. You ask the database for every instance where two or more photos have the same coordinates. There are, it might turn out, only two. Looking at the photos, you see that one is a gas station, while the other is a Starbucks. Bingo.



LOST IN SPACE: The GPS coordinates captured by MyLifeBits can be strung together to show a journey, using an ordinary Web-based map. Since each point on the map corresponds to a photo taken by a SenseCam, the photos can be displayed along with the map, or, as here, a variety of trips can each be represented by a single thumbnail. At the lower right, the start and end times of each trip are also shown and can be scrolled through.

ity in a day—it can hold 2000 low-resolution shots. Playing the images at a languid 6 images per second, you get something that looks like a 1905 newsreel, hardly better than the tiny books of flip pages that come in boxes of Cracker Jack. But for reviewing your day, it's brilliant. At that rate, it would take Mr. and Mrs. S. only 3 or 4 minutes to run through a 12-hour day. In practice, they stop at key frames and discuss them in detail. According to Wood, Mr. S. says these sessions have made a "dramatic difference" in his wife's ability to recall things.

MICROSOFT KNOWS it has just scratched the surface when it comes to finding ways to use a SenseCam, so in July, it put out an open-ended request for proposals. Gemmill says a total of

come up with made the data look like a spreadsheet. Each row represented a document, and each column some fact about the document—the document type, the date it was created, a GPS coordinate if there was one, and so on.

So a Microsoft research group in Redmond, called Visualization and Interaction for Business and Entertainment, is creating ways to take the people, places, and events hidden inside the MyLifeBits database and show them onscreen. In one interface, for example, an object, say a sales meeting, shows up as a tinted oval. When you zoom in on the shape, it expands into a field of dots, like a Seurat painting dilating into points of color as you walk close up to it. Each dot represents a document about the meeting—e-mails with the agenda, PowerPoint presentations that were given there, photos,

and so on. Dots that have other things in common—perhaps they're presentations that were given the same afternoon—are clustered together. Clicking on a dot opens the document it represents. The group is experimenting with ways to use the size and tint of the dots to convey even more information. It will be a powerful way to look through your life bits.

Suppose you want to search for a specific fact. What was that delicious cabernet that was served at the sales meeting? The visualization and interaction team is developing software to help you find it. The software figures out what you're looking for and automatically produces documents that might be relevant. So as you're searching through your SenseCam photos for the wine label, for example, a new window might pop up onscreen with an e-mail message, from the week before the sales meeting, with the word "wine" in it. You completely forgot that the dinner menu was sent around in advance! The software, in effect, is looking through the tens, soon to be hundreds, of gigabytes of your life bits for you, without your even knowing it.



AHEAD OF HIS TIME: In a 1945 article, presidential science advisor Vannevar Bush predicted that one day it would be possible to record all the world's facts and store them in a Web-like database.

STORAGE IS CHEAP, and getting cheaper and cheaper, at a rate that exceeds the Moore's Law formula.

It's been doing so for a long time. "The *Encyclopedia Britannica* could be reduced to the volume of a matchbox," Bush predicted in that 1945 magazine article. At the time, a complete set, from A to Z, occupied a shelf a meter long.

A gigabyte's worth of hard disk space cost \$1000 in 1992 but only \$9 ten years later, according to a recent study by Steve Gilheany of Archive Builders, a Manhattan Beach, Calif., consultancy. Gilheany believes there will be similar declines for 1024 gigabytes, that is, a terabyte: about \$2200 today, one tenth that in five years, and a minuscule \$20 in five more.

A terabyte would hold about 900 sets of the *Encyclopedia Britannica*. After all, it's almost all text, which has little of the bulk of photos or audio files. But even those media are easily archived nowadays. "Every bit of sound you ever hear and everything you ever say will fit into a terabyte," Gemmell says. "And if you're putting low-resolution pictures onto a terabyte disk for an entire lifetime, you could take one every 3 minutes, from birth on."

How much of your digital doppelgänger can fit into one terabyte? Gemmell says that Bell's life bits have been growing at a steady 1 gigabyte per month. At that rate, the average person wouldn't start the second terabyte until the age of 84. That's 13 years away for Bell, but he keeps finding new things to add to his life bits. Back in January, he went to his doctor for a detailed physical exam. A few days later, he got a 30-page report, nearly incomprehensible to the layperson. Naturally, it was turned into a PDF file and dumped in with the rest of his life bits. "It had stacks of genomic data," Bell says, "and it was in a nice format and all, and it had all these medical terms. I didn't understand a word of it. So I had an hour-long phone call with my doctor, and we went through it practically gene by gene. I couldn't possibly have taken notes. But I happened to be in my office, where all my calls are recorded."

Bell has been talking to lawyers and physicians recently about how critical it is to collect audio. "Of course, there's all kinds of privacy problems, and all," he says. "But the fact is, for the patients or clients, it's really crucial, because you can't remember this stuff."

The next version of the SenseCam, being designed in Cambridge, will record sound from the moment it's turned on in the morning. It will dump it all to the PC—the personal mainframe—every night, along with all the other data it captures: temperature, tilt, photos, and GPS coordinates. Future versions may try to capture heart rate and other biomedical data, its designers say. Why not?

Privacy is one reason. It's not a concern yet, because your life bits are your own, and no one has access to them except you, or a trusted loved one, such as Mrs. S's husband. Bell's life bits cascade through a number of different hands within Microsoft's own research labs, but when I asked Gemmell if the request-for-proposal winners would get a copy of the database, he blanched at the thought. "There's phone calls and receipts and e-mail in there, all kinds of personal stuff."

Yet we can imagine the day when your life bits show up in court—first, voluntarily, perhaps to validate an

alibi, later by court order if, say, you're a material witness to a crime. Eventually, this may be the way we settle ordinary disputes, from legal contracts to your daughter's latest tantrum. "You promised we'd go to Disneyland!" she screams. "I made no such promise," you say. Did you? Let's take a look.

Earlier this year, in the journal *IEEE Multimedia*, Frank Nack, a researcher at the Center for Mathematics and Computer Science in Amsterdam, the Netherlands, asked whether we always want to have our life bits around to answer such questions.

"Forgetting makes our life bearable," Nack wrote, "and is closely related to cultural concepts, such as forgiveness and absolution. I wonder if removing this human imperfection, namely to forget, would do more harm than good." Will we create a society of conformists, he wondered. "If life logging was the norm, how would we break social rules and norms?" Would we be as creative? As free?

We'll find out, soon enough. By 2015, all your life bits will fit into two or three matchboxes, about 50 bucks' worth of data. Your smartphone-sensecam will dangle casually around your neck, snapping away. Want to know what you wore on that blind date you went on last month? How many glasses of chardonnay you drank, which Web pages you viewed the next morning, whom you called on the phone and what you talked about? Where you were for every minute that week-end? Let's take a look.

TO PROBE FURTHER

Gordon Bell's home page is <http://research.microsoft.com/users/Gbell>; Jim Gemmell's is <http://research.microsoft.com/~Jgemmell>; and Lyndsay Williams's is <http://research.microsoft.com/users/lyn>.

The SenseCam is described in more detail at http://research.microsoft.com/sendev/project_sensecam.aspx.