

Still Seeking: Signposts of Things to Come

Glorianna
Davenport
MIT Media Lab

Where is Lewis Carroll when we need him? What delectable experiences will a storyteller on a par with Antonioni, Spielberg, Kubrick, or Buñuel offer us 20 years hence?

Last year in this column, I plainly laid out my prejudices. I seek dynamic, adaptive story environments. The roots of my quest are both personal and political: personal, in that I enjoy my kind of good story; and political, in that to communicate with a diverse audience, a good story should satisfactorily support readings from several different points of view.

This year, the buzz about standards, compression, and the “interactive set-top box” has moved out of the public eye into quieter back rooms. It has largely been replaced by news about who is making what for publication on the World Wide Web. The good news is that we now have a primitive infrastructure capable of delivering most of the multimedia objects and functions we once thought essential. The bad news is that this infrastructure is currently underpowered and extremely cumbersome to use.

As a desktop publishing medium, the World Wide Web harkens to the old model of pamphleteering, where anyone has the right to advertise opinion without censorship or editorial interference. How far can we go down this aesthetically and sociologically independent path before regulation and other forces converge to tame it, imposing a layer of moderation or editorial control between author and audience?

It is difficult to do justice to recent phenomena which, this year, are leading us to dynamic, adaptable story environments. The reality is far different than what we were preparing for just a couple of years ago. Between network software and hardware, e-mail applications, and the World Wide Web, we are rapidly becoming a wired nation. How much experience can we amass before technical limitations become barricades to

invention, or before the Web becomes so familiar—or so tamed—that we once again long for something new?

Electronic babysitter

If we look at the most banal but profound reason for structured media—to communicate interactively with other people—e-mail stands out as a remarkable evolutionary achievement. As an asynchronous messaging system, it is by nature personalizable. Although e-mail is unidirectional and does not occur in real time, it enhances communication without disrupting concentration.

E-mail has certain not-to-be-forgotten parallels with the telephone. Shortly after Alexander Graham Bell invented his telephone, a government official came to look at it. As the apocryphal tale goes, the official remarked, “Very nice, very nice, there should be one in every village.” In those pre-enthusiast days, no one envisioned a time when parents would need to threaten installation of a pay phone in the house if their 12-year-old daughter could not restrict herself to conversations under two minutes. Today, for an extra fee, the telephone company offers its customers various forms of “call blocking.”

Over the last few years, e-mail applications have spread from computer enthusiasts to the masses. I routinely find myself corresponding with family members and friends who have just joined the ranks of the e-mail literati. With the explosion of on-line services available to anyone with the money, people from all walks of life are coming within the reach of e-mail.

And what of video mail? What has become of the dream that video mail would be the “killer” application that would drive higher-end PC’s into the office? Two years ago, this seemed reasonable; today, I wouldn’t bet on it! In fact, things appear to be going precisely in the opposite direction—higher-end PCs are flooding our homes—and, as

Figure 1. The cast of the Spot, a Web-based soap opera.



Participatory television

This year, as I continued my quest for the dynamic, adaptable story, I occasionally wondered what had happened to the much-ballyhooed industry thrust toward interactive television. In general, discussions about standards (such as MPEG-4) became surprisingly low-key. News of interactive television trials were barely discussed in the press. Perhaps Barry Diller took the excitement out of home shopping by implementing it, showing us how the reality compared to the theory. Even worse, video-on-demand, the other supposedly “killer” application—lacking pause, fast-forward, and other popular VCR functions—appeared to be a service still in search of an actual customer. Even talk about the set-top box seemed muted. I only hope that the squabbling over the design and capacity of these special-purpose computers will last long enough to permit an upgraded general-purpose PC to fill that role. These observations might just reflect a particular filter on my information, but, having reviewed a wide variety of literature, I believe that we have come to an interesting crossroads.

In the future, “television” promises to be digital and interactive. However, for the computer to replace the television set—or to become embedded within its circuitry—we need to establish an infrastructure that uses robust construction tools to manage extensible content. In recent years, the development of interactive systems has split between the delivery of traditional entertainment (controlled switching to tap into pre-edited streams of information) and education (flexible hypertext-like surfing through information spaces). We seem to be reaching an apogee on the hypertext side of the equation. However, some recent activities, particularly the World Wide Web, blur this distinction.

The Web is a publishing medium that, in theory, takes graphical exposition to heart. Today, it fills the role of a modern bazaar, and everyone has their own flag flying. For example, the virtual gallery put together on the Web by the Interactive Media Festival, held at the Arc Gallery in Los Angeles but available world-wide at <http://www.spark.com/>.

However, a virtual gallery is not a story. So where to next? How about visiting the Spot (<http://www.thespot.com/>), a soap opera produced by Fattal & Collins, an advertising agency, and Prophecy Entertainment (Figure 1). It expands daily with new episodes and encourages voyeuristic viewers to follow the lives of its characters in full-color photos, text, and video. While for now

usual, our children are leading the way.

Almost half a century ago, television suddenly became America’s premier babysitting service, while the common social experience of news, soaps, and sports became required viewing for most adults. Sometime in the 1980s, television as babysitter was replaced by the new \$\$\$ toy on the block as we entered the heyday of the Atari, Nintendo, and Sega video-game empires. For many of today’s kids, CD-ROM-based games and adventures on the Internet—experiencing them and making them—have become paramount.

The rise of computer sales for home use should not come as any surprise. Most parents would like their personal computer to remain personal, which means that (for those who can afford it) a second home computer has become a necessity. Frequently, the kids’ computer is better than the one the parents control, loaded with “educational” features. I hear a constant stream of stories from proud parents whose son or daughter has mastered the mechanics of their machine: sound effects, PICT backgrounds, network access. In fact, the younger generation participating in the on-line community is an increasing reality of family life. We might well guess that so long as this personal-computer forum remains active, computers will be a more popular gift for the playroom than another television set. Will there come a day in which the World Wide Web or its successor becomes the nation’s babysitter of choice?

it's subsidized by Fattal & Collins, the Spotmakers plan to include advertising eventually.

The Web even gives a new look to prime-time TV. Transcoded directly from television, NYPD Blue Comics! (<http://www.media.mit.edu/people/mmassey/>), developed by Michael Massey, combines SalientStills images generated from key video scenes, complete with dialogue balloons. And consider "Lurker" (<http://lurker.www.media.mit.edu/registration/>), a recent Interactive Cinema "Thinkie" that not only uses text and posted video segments to tell a story (Figure 2), but also enjoins the audience to exchange e-mail with characters and each other as they help the fictional hackers solve a life-or-death mystery. Figure 3 shows the registration screen from "Lurker."

The coincidence of three attributes makes the Web significantly different than anything that has come before:

1. A huge body of posted information has already accumulated in the system, and it is constantly growing. At first, most of this posted information was text, but increasingly the Web is a playground of stories built in many media types.
2. Virtually anyone who wants to advertise objects or ideas can set up a small server and do so. This makes the landscape both more democratic and more constructivist than any prior large system. Anyone can be a player, a designer, a communicator, a seeker!
3. Finally, certain Web browsers and helper applications support a full range of media—still pictures, sound, video, and active programs—as well as graceful degradation when different levels of success are encountered. The Web is therefore a potential testbed for new ideas about interactive television.

As it now stands, the World Wide Web is inadequate to totally fulfill the promise of multimedia, which requires that a shape be given to the rabbit hole so that anyone who falls down it will have a magical experience. The Web environment must overcome several shortcomings before that can come true:

- As a browsing environment, the Web does support extensible content, although the approach to annotation and searching is crude, cumbersome, and not particularly effective.

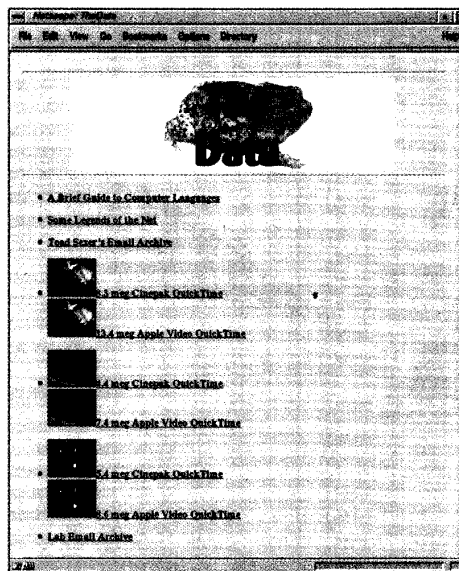


Figure 2. The Media Lab Interactive Cinema "Thinkie" called Lurker involves participants with a mystery using story, video, and interactive e-mail.

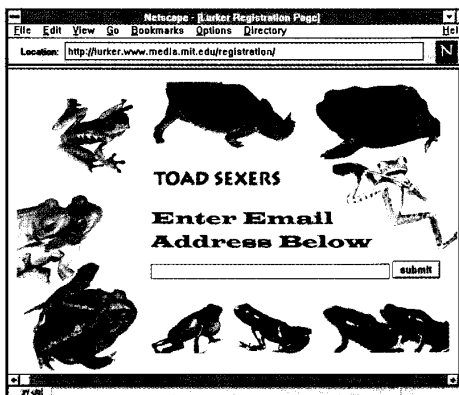
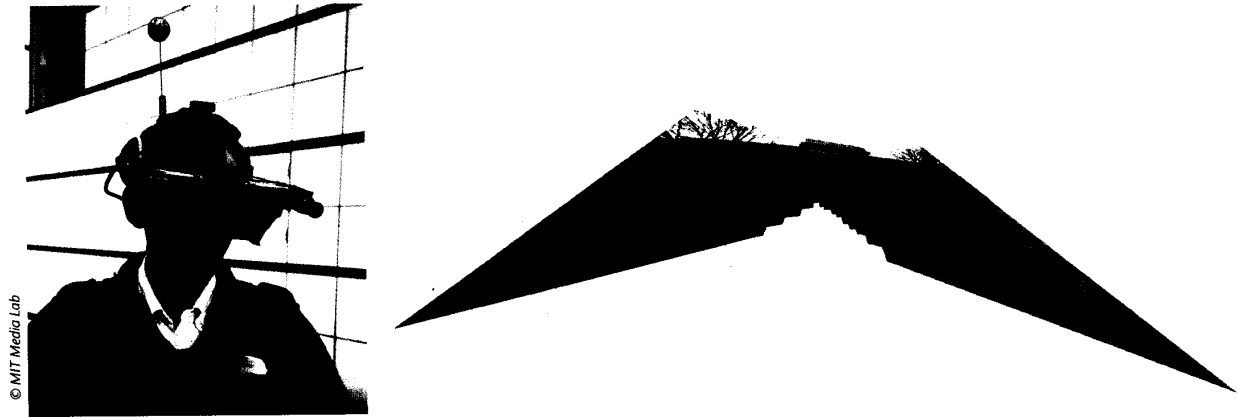


Figure 3. Lurker visitors register to join in the mystery-solving.

- Likewise, the constructivist environment remains crude. While many contemporary word processors, layout programs, and special languages (such as Sun Microsystems' Hot Java) format to HTML, writing in HTML remains tedious and not at all dynamic.
- Today, the aesthetics of the Web are ultimately driven by the constraints of bandwidth: In universities and corporations possessing high-end equipment, graphics and motion video segments can be downloaded in a reasonable (but certainly not instant) time frame. For the average consumer at home, even a high-performance modem provides an intolerably long downloading time.



© MIT Media Lab

Figure 4. In the Wearable Wireless Webcam experiment, Steve Mann walks around the MIT campus with a camera strapped to his helmet (left) transmitting what he sees (right) over the Internet.

■ Current browsers simply do not do enough with fonts, color, smart downloading of movies and sound, layout, or 3D.

The Web lacks the ability to display immersive temporal stories that are seductively visual. On the Web, text rules. This happened largely because the granularity of word/sentence/paragraph and the structures of chapter, table of contents, and index are historical primitives that allow you to drill down into, as well as navigate through, textual structures. Text operates with a limited set of standardized characters, each with its own standardized digital representation, and in that way ideally suits pattern matching in a computer search. Text tends to be stable and relatively invariant over time.

On the other hand, a stream of moving video imagery is highly variant over time and appears almost monolithic in its structure. Indexing is just as fundamental for movies as for text sources, yet video indexing involves transcoding both the time base and the associative values within the imagery into text. Although video flows from one granular piece to another—shot to shot, scene to scene—the granular elements within a frame are difficult to isolate and describe. There is always the question of what is important and why. Trying to use image processing to tease out an Edit Decision List because producers are afraid of including it makes the problem even more difficult.

Nonetheless, activity on the Web is remarkable and noteworthy because, even in its crude contemporary form, it puts into play a more salient vision for new media than anything that has come before. Unlike the Warner-Amex QUBE experiments of the early 1980s or the current

visions of video-on-demand services, the Web is democratic and constructivist by nature; it lets everyone in to play. In this sense, it is an open authorial architecture, both extensible and interoperable. Its hypermedia-like structure allows us to do things substantially different from traditional stream-based television.

Accessing the Web is like having hundreds of television channels on which anyone can broadcast. However, making a “best-of” list of the Web does not do it justice. The medium does not cater to traditional forms: it is something new, a form that mixes conversation with story.

Much of what I see is goofy stuff, but in many cases I get a glimmer of ways these experiments are modifying our idea of “interactive television.” Not surprisingly, I find the most interesting Web pages—those I learn the most from—tantalize me with quasidynamic construction of content. In general, these constructions are serious and playful at the same time, such as the documents generated with the Wearable Wireless Webcam, a head-mounted video camera uplinked to the Internet. In this nomadic application, the “author” exploits his right to be a walking, low-power TV transmitting station (Figure 4). However, it is important to remember that broadcasting a stream of incidental images is not the same thing as telling a story. (To get to the Webcam page, follow “http://media.mit.edu” to the MIT Media Lab Home Page, then select “Steve Mann”.)

Progress can never be measured by one invention or one trend. Tapping the potential of any interactive scenario adds to our ability to appreciate the ultimate transformational environment when offered.

Marking a trail

Without grinding through a lot of old history, Vannevar Bush's vision of Hypertext (as described in his landmark 1945 article "As We May Think") included the notion of "memory traces." As a researcher traverses the literature in the library, she can leave a trail of electronic "footprints" that others may follow—a trail that might include personal commentary and other enhancements of the material. In many ways, an individual researcher acts like an editor: evaluating the relative value of materials, discarding the irrelevant, collecting and connecting the relevant, and attaching informed commentary to selected materials.

Today, this vision has been augmented by the programming of "social agents" that can search for similarity between large feature sets and generate on-the-fly a composite path, based on the exploratory roamings of many individuals. Agent makers might choose to extract feature sets that reflect a specific special-interest community, as in HOMR: Helpful On-Line Music Recommendations (<http://ringo.media.mit.edu/ringo/ringo.html>). Under this approach, trends discovered within a community (Figure 5) are shaped into a pervasive editorial voice, which is then used to selectively steer individuals through the larger information space (Figure 6). As the source and scope of the editorial voice change, the development of social and intellectual groups might also change.

Vannevar Bush recognized the power of leaving a trail of bread crumbs through expressions of invention and opinion harnessed by editorial functions. As viewers, we can only reconstruct the story when the story resembles a familiar place. But how familiar must it be? This year, I have met many writers of television and movie fame who would like to work in the new media but are perplexed by the lack of a formula. At the same conferences, researchers hang out with artists, hoping to extract some expert knowledge: the structural and procedural behavior of storytellers. Slowly, the big conferences—CHI, Siggraph, Digital World, Multimedia—are catching on. In the midst of this maelstrom, I must offer a word of caution. The medium is new; it has a right and a privilege to be other than what we have known in the past. Re-read Vannevar Bush.

MM

Contact *Visions and Views* editor Glorianna Davenport at Media Arts & Sciences, Massachusetts Institute of Technology, 20 Ames St., Cambridge, MA 02139, e-mail gjd@media.mit.edu.

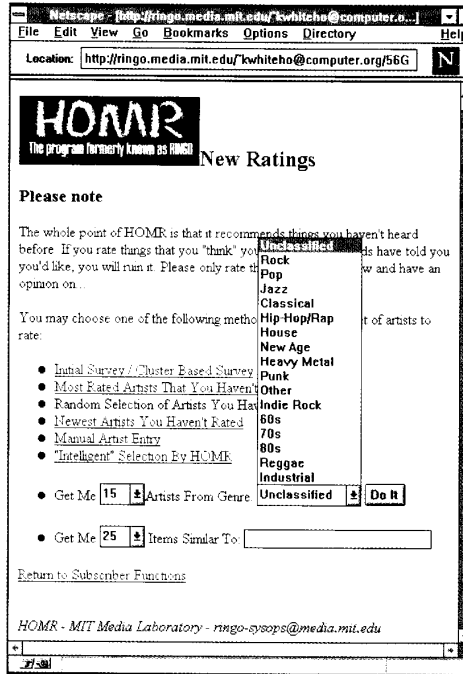


Figure 5. New members of the Helpful On-Line Music Recommendations community create a profile by rating musical acts in a number of categories.

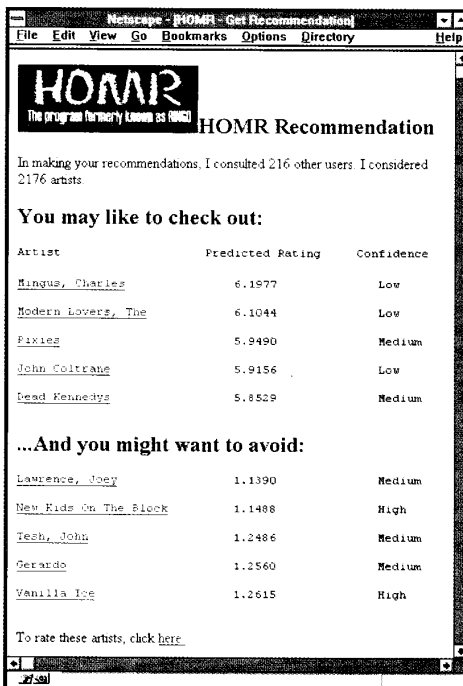


Figure 6. Once HOMR receives a list of user ratings, it recommends acts to pursue and avoid, based on similar users' recommendations and powered by the Ringo++ language.

Fall 1995