What is the shape of things to come?

Paper collectors all know the thrill of a surprising discovery. Odd variations of well-known printed matter show us the idiosyncrasies of the toy train development process and provide a perspective on America in an age very different from ours. Many of us are now exploring toy-train-related Internet sites to learn about today's toy train development process. While the Internet brings a huge volume of information to our fingertips, you may be surprised to learn just how ephemeral that information is. When web pages are changed or removed, most of the time they are simply deleted and lost forever. Today it is fairly easy to lay your hands on an original 50-year-old Lionel consumer catalogue but impossible to view Lionel's early web pages - Lionel isn't saving them!

A casual e-mail survey of toy train websites that would be fun to view again in five, or ten, or fifty years revealed that only the Showcase Line and K-Line currently archive their pages. A staffer at *Classic Toy Trains* magazine summed up the situation this way, "I don't believe anyone has saved anything, beyond, perhaps, a few images. I think this is primarily because it is 'old news' and nobody wants to waste the server space storing obsolete information."

That statement should send cold chills down the spine of any collector. So, what will a "paper" collectors' meet look like in 2050? Yes, the industry is still printing tons of paper, but what can we do about all this electronic material? There aren't any easy answers, but there are several possibilities to consider.

Possibility number one: Someone out there is saving web history already. Yes, web history is being saved at the Internet Archive, and you can visit their web site—archive.org to learn what they are doing. Since 1996, this non-profit organization has employed web robots to download 1 billion pages (text only) from the Internet. "In the past four months we have begun images crawls (automated data collecting) and have about 1 Terabyte. We expected a larger collection, but a large number of URLs have robot exclusions and cannot be crawled," according to Brewster Kahle, founder of the Internet

Archive. For the past year, they have also been archiving pictures at a rate of 200 per minute. In four years, their colossal undertaking has sucked up 14 Terabytes of web data. In round numbers, a terabyte is 1000 Gigabytes. To get a sense of scale, the Library of Congress' collection of 20 million books (not counting pictures) totals roughly 20 Terabytes. Kahle's archive should be colossal by the year 2050.

In a recent interview with CNN, Kahle compared his work to saving old newspapers, "If we had been selective, we probably would have kept all the articles and thrown away those ads, but it's the ads that the historians really like. That's what gives them a much better glimpse of what life was like." . . . Ain't it the truth? The downside of the Internet Archive? For now, you have to learn Unix programming to search their collection. Still, what we love must be in there, somewhere.

Possibility number two: Print out the website in hardcopy. This is a comparatively low-tech solution. Color printers aren't perfect, and who knows how durable the images are, but you will have something to hold in your hand and look at. Remember that even if you archive an entire web site, you have only saved page images and not the fully interactive site - not unlike photocopying a paper catalog.

Possibility number three: Toy train paper collectors can begin archiving web pages themselves. It is now possible for the home computer user to buy CD recorders and create ("burn") their own CDs. CD is short for "CD-ROM" and means "Compact Disc-Read Only Memory". A CD-ROM is physically identical to an audio Compact Disc used in a CD player, but the bits recorded on it are interpreted as computer data instead of music. A CD-ROM can store large amounts of data—up to 1 Gigabyte, although the most common size is 650 Megabytes. A single CD-ROM has the storage capacity of 700 floppy disks, enough memory to begin your archive of toy-train-related web pages.

We asked veteran Silicon Valley programmer Chris Keith for advice. He said, "Archiving a web site can be a complex undertaking because the site can be a lot more than just a bunch of HTML pages (Hypertext Markup Language - HTML - is the computer language used to create documents on the World Wide Web). It can include lots of code, data stored in a database and so forth. What kind of stuff do you want to archive? Parts lists and descriptions? How-to articles? The simplest (though most tedious) method is to go to every page you want to archive and save it to your local disk, then collect them all on a CD. The links from each page may go dead later, but you will have the raw data. This would work for HTML, but I don't know how it would work for pictures—you would probably have to go through yet another separate step to save the picture files."

CD technology has made it possible for us to collect toy-train-related web pages, though as Chris pointed out, there are still some

TRAIN WEB SITES

by Greg Johnson

open questions concerning the best way to do it. And aren't there some problems with CDs? Yes. The two biggest concerns are CD shelf life and rapid changes in CD technology itself. A CD is a thin aluminum layer sandwiched between two pieces of plastic. Laser light transforms electronic data into microscopic pits in the aluminum. There is a concern that air can infiltrate the CD sandwich and destroy the data by oxidation. The smart money says that the average CD will last for a human lifetime. The key word here is average. Next, the standards and technology of CDs are changing rapidly. It is probable that the encoding that converts electronic data into laser-burned pits on a CD will be very different in the year 2050. The information will be packed ever tighter; the CDs themselves will be smaller. Perhaps the basic technology will be scrapped altogether. Do you upgrade your entire library of archived web pages each time the "latest and greatest" hits the market? Alternatively, do you maintain a museum of antique computers to ensure that your data can be read? *When was the last time you saw a roundhouse in use?*

And how about copyright? Popular press stories about downloading royalty-free music through the Napster web site have brought the legal phrase "fair use" into everyday conversation. Is it legal to record commercial web pages for personal use? Could you sell 1990s toy train web pages at a York meet in 2050? Paper catalogues were meant to be given away. What are the legal issues surrounding the "ownership" of electronic data? The legal community will be wrestling with these questions for years to come. I have come to believe that toy train manufacturers may already recognize the potential profit in selling their electronic websites and catalogs on CD. CD duplication costs are negligible and electronic toy train manufacturers must recognize that the Internet has brought about a fundamental change in the traditional ways of doing business. What are we willing to buy from them? Many toy train makers have already begun charging for their paper catalog, can their electronic data be far behind?

I hope this article has provided you with a small preview of the collector world of 2050 when the TTP&M becomes the TTM&M - that is the Toy Train Media and Memorabilia Group! ∎

Further reading:
A Brief History of the Internet:
http://www.isoc.org/internet-history/brief.html
Archiving the Internet: A Brief History:
http://www.ils.unc.edu/~martkr/archives.html
Discussion of recording web pages on CDs at Deja News:
http://www.deja.com/home_ps.shtml
CD Recordable FAQ: *http://www.fadden.com/cdrfaq/*
Mass storage information:
http://www.zdnet.com/pcmag/stories/columnists/machrone/0,5655,2245813,00.html

TTP&Mer Greg Johnson - he's an Information Technology manager at a major university - combines a knowledge of the latest technology with a strong interest in history, and has written on both topics. We thank him for this article demonstrating that there's always something new in train collecting.