ILLUSTRATIONS IN REAL WORLD ELECTRONIC PUBLISHING

Mr. Alfred L. Fenaughty
President
Information International Inc.

The emphasis of my remarks this morning is on "real world" because that is where all of us must function in the real world. And it is in that world of too little time, too many demands and too few staff that we must bring technology to our aid. Technology must add to our resources and in our particular "real world" it must assist the process of printing and publishing.

The ads and the air waves push "desk top publishing" and "personal publishing systems" as a complete solution. Are they solutions in themselves or a tool to attack an annoying bottleneck in real world publishing?

Since virtually every publishing enterprise is driven by deadlines, it is not surprising that we spend our resources looking for ways to beat the deadline. It is routine to speculate....Can a document be produced faster without quality degradation? Can we improve upon the quality; or revise more frequently ; or produce more versions of a document and still get it out on time.

The task faced by publishers is daunting. Who even thinks of how yellow pages, parts manuals or even encyclopedias are put together, let alone where all of the pictures come from....or more to the point....where they are stored.

How many repackagings of the same data base are needed to provide quick access to the right information for the plane you last rode in or to maintain your car.

The volume and complexity of routinely published documents in every field is staggering: aerospace; government manuals; military equipment and vehicles; farm implements; consumer appliances and on and on. We aren't talking "one per" here either. These items routinely require parts catalogs, as well as instruction manuals on how to use and how to maintain.

The purpose of these manuals, of course, is to educate, to inform and to instruct. We all know the best way of making a point. Not with words only, but with pictures, illustrations, and diagrams.

A publisher's nightmare? Not really. Just a large real world problem for which suppliers like triple-I develop solutions. Automation of this process has been a long time in development.

The electronic impact on publishing started with teletype driven line-casting machines and moved on to photo typesetters.

In the late 1960s, RCA's mainframe computer brought batch page composition to the process. This technique of standardized full page composition using page format commands remains in use at such large scale operations as the State of California, publishers of all its printed material from daily legislative documentation to election material, technical reports and even of all the state's high school and college text books.

The technology advances introduced by RCA at this time included:

1. Single keystroke capture which allowed editorial keyboard input to go directly to composition.

2. Full page, rather than galley, phototypesetting.

The promise of a major market prompted many computer companies to enter the field of electronic publishing. As with most emerging industries, a shake out saw many enter and quickly leave....IBM among them. But others made a lasting impact.

The advent of full page typesetters set the direction toward what has become today's electronic publishing systems. Early RCA systems served as prototypes for today's standard four-step pagination flow chart:

1. Input terminals for text

2. Computers to manage the data base

3. Composition to arrange text in page form

4. Typesetters to record complete pages directly to film or photo paper.

Many large scale users applied this model to their publishing operations. Even Federal and state governments saw their way clear to automating the production of voluminous documents and the ubiquitous tax form.

The 1970s brought marked change to the publishing industry. The changes centered on cost, performance, size and speed. Minicomputers improved the price/performance ratio. Computers cost less and did more.

Users welcomed such performance enhancements as:

1. "Exception" dictionaries to handle errors due to hyphenation logic;

To answer, we must examine the production requirements of today's publishing systems. To that end, I am presenting the results of an informal survey of our customers. Although it is incomplete....some interesting applications are classified and, of course, I had no access to users of competitive systems....this sampling will give you a feel for the requirements in commercial and technical publishing.

Our equipment has been used to produce most of the world's encyclopedias, but in earlier days, the illustrations were pasted in. The French Univeralis was one of the first encyclopedias fully electronically composed with integrated end-to-end composition and all illustrations in place. About eight years ago, the complete set of volumes was totally electronically produced.

Tabular work, such as that required in financial documents and annual reports, is also easily accomplished with electronic composition.

Interestingly, despite the growing availability of CAD/CAM, technical illustration is still dependent largely on hand-made drawings produced by draftsmen or illustrators. This art is captured as digital facsimiles with compression to reduce storage and transmission cost. Vector data is often used for wiring diagrams because it permits the recording of a higher density of information per square inch and requires minimum data.

Photographs are heavily used in applications where documentation must be understood by operators and mechanics with limited formal education e.g. where the documentation is used to guide assembly line workers.

Clearly, electronic publishing becomes a massive data management task with the addition of ILLUSTRATIONS. In comparison, text becomes almost insignificant in its relative requirements for access. As this survey shows, even data compressed line art requires substantial digital storage, on the order of 100 kilobytes per drawing.

New illustrations are added constantly; at the same time old illustrations are updated or purged. The data base is in constant flux...and growth. Such incremental expansion disguises a problem. Additions and revisions to the data base are gradual. However, at publication time, large volumes of this data must be rapidly accessed. This demand strains most computer systems.

The operating systems and the file structures of most computers, as well as their basic architectures, are not designed for high speed processing of large numbers of very large files. Ethernet is a perfectly adequate networking system in a low volume text application; it collapses with only the smallest amount of illustration data.

In addition, operating systems such as UNIX do not, in their standard form, operate in a real time environment. They neither provide automatic backup of data, nor easy hardware reconfiguration in the event of sub-system failure. A deadline oriented, high-volume production system operating in the real world MUST have these safeguards.

To date, the solution has been for each supplier of high production publishing systems to develop its own or to substantially enhance standard operating systems. Off-the-shelf solutions simply are not adequate to this level of performance.

Turning our attention now to scanning requirements in high-volume applications, the scanner must automatically capture the data in the form required for subsequent publication. Time and cost considerations prohibit operator interaction to establish gray-scale threshold levels or to crop, size, position and de-skew individual pieces of art.

Because of the volume of new illustrations, many aerospace applications depend on a microfilm intermediate for illustration input. With this type of scanner, rotational squareness, sizing and position reference coordinates are assured in the planetary camera operation. Pin registration of the film in the scanner retains positional reference. Dynamic thresholding automatically compensates the B/W threshold to assure uniform capture of fine line work. Software also deletes blank border areas while retaining the input coordinate reference, thus minimizing data storage and reducing recording time.

Most desktop scanners lack the capability of assuring an exact location or the rotational alignment of the digital image. They provide no threshold compensation for the different contrast ratios of pencil, India inked or printed input copy. In almost all cases, the optical flare assures that there is no single electronic threshold level which will capture fine line work (thin black lines and thin white lines) at both the edges and the center of the picture.

In triple-I's direct reading scanners, we find it cost effective to provide graphic arts pin registration systems to position the copy relative to the sensor. We also provide sophisticated retroreflective optical sensing systems with a large dynamic range and two-dimensional image processing hardware to assure capture of the image data from the wide variety of tech doc illustration sources.

Today, such scanners sell for over $100,000 and are produced in limited quantities. As the market recognizes the need for positional, rotational and gray-scale quality, tooling for higher production quantities thus reducing cost will be justified.

2. The ability to move text between files.

In the 80's we found further enhancements such as:

1. Full hyphenation dictionaries with quality-graded hyphens;

2. Full spelling lexicon with immediate user access and the availability of even more computer capacity at lower prices.

Despite large systems vendor deprecation of "desktop" solutions, most are beginning to capitalize on the performance advantages of PCs and other standard work stations and are integrating them into their own larger scale product.

For instance, triple-I's front-end editorial system, TECS/2 by Morris, includes a 120,000 word dictionary which provides graded hyphenations and also an instantaneous spelling check; and a split screen display which makes it easy for writers to move or swap text within a file or between files by word, sentence or paragraph.

Nothing particularly unusual. These are typical features of any modern editing terminal EXCEPT....in this case the entire software package....ALL of the capability....the text stream manipulators; the composition; indices to file; even the complete spelling and hyphenation dictionary is duplicated on EACH one of the work stations. Although with networking the terminal cost is today higher than dumb terminals tied to a mainframe....its blinding speed....its enhancement of user productivity....and the elimination of the mainframe....makes it most cost effective in production environments.

Integration of illustrations was first used in production by aerospatiale in 1972 for the SST Concorde.

In the '70s direct microform recording required publication of illustrated documentation in page order sequence demanding more comprehensive composition systems and introduced automatic index generation.

The need existed in many applications to record true page size films as well as standard microforms. Because of its unique capabilities to record typset pages on all film sizes, the COMp 80/3 has become the standard for both commercial and military aerospace documentation. The initial phase of electronically integrated technical document illustration was limited to simple line art, usually hand drawn isometric views and piping or flow diagrams....with vector plots often used for wiring diagrams.

That all changed in 1976. Weekly news magazines needed to rapidly transmit high-quality (symbolic) plate-ready pages to multiple printing sites across the country. This led to the integration of symbolic electronic halftones....initially black and white, later full process color.

Now, continuous tone photographs could be electronicall cropped and scaled to final output size; then screened to a specified halftone screen angle and mesh....and the data compressed for storage and transmission. This became the working model for all the major news weeklies TIME, NEWSWEEK, U.S. NEWS and WORLD REPORT and most recently, THE ECONOMIST. Today, over 30(magazines are produced using variations of this first system.

The same basic technology applies to technica documentation, although in this application the demand fo color, for the moment, is limited to a single color layer o spot color to highlight a wiring or piping path; not full, fou process color.

A decade ago the cost of electronic publishing equipmen limited applications to either:

1. Very high page VOLUME requirements, or

2. To situations where TIMELINESS was crucial enough to justify the investment.

But with time technical advances such as higher density memory chips, bit map WYSIWIG displays and low cos laser printers have substantially reduced equipment cost and the need for typographic expertise. Other applicatior benefits are evident. The dry print process of a laser printei is suitable to office environments. This provides the opportunity to prepare page proofs....and in some instances copy for final distribution....at the site of authorship.

Higher resolution screen displays, higher quality laser recorders and more reliable keyboards are all available at more reasonable prices than previously thought possible while the substantial computing capacity of the PC has allowed the editing and composition work load to be offloaded from the mainframe computer to the termina itself.

Smalltalk-based software has reduced the computer jargon....a major obstacle for new users.

So how do all of these advances in personal computing fit in electronic publishing? What is left for the large system to do now that every "A type" executive has computing power on his desktop?

Obviously, high page volume publications require a number of copywriters, engineers and editors operating on a common data base. PCs potentially offer substantial power to each user.

Unfortunately, networking PCs to share a large data base is still a development challenge. With minicomputer central processors the simple task of determining who needs service and for what purpose utilizes a significant portion of the processor's horsepower. There has been little progress in improving this lack of efficiency.

Triple-I's answer is to distribute the data base, letting each PC handle its own files on its local hard disk and connecting all PCs (and the port to the pagination system) in an enhanced token ring network. Use of the local disk for each user's working storage keeps most of the traffic off the network.

When needed, a central index of all files assures access to any file. When one work station's file is completed, the file is transferred to the work station next up in the editorial hierarchy.

For illustrations, our approach is to capture the data at the scanner in final publication form. We store the data in duplicate, but put it away in its final resting place at scan time. From then on, it is referred to as needed at that location and is transferred for recording only for the final output page.

All references to that scanned image for interactive composition and for soft proofing require only a low resolution version of the image. This image is also created as part of the scanner process and stored in the location as the high resolution image.

Keeping the number of transfers across the network to a minimum, while assuring appropriate file security with built-in file redundancy and self-purging, has proven highly effective. It provides rapid response and makes effective use of the new technology. Of course, the MIS manager whose comfort level is dependent upon the whirring of "HIS" disks, will have to make a psychological adjustment....or come up with a better system to improve the access time to "HIS" centralized files.

SUMMARY AND CONCLUSION

Today we have reviewed our progress from hot metal to photographic methods, to today's digitally-generated output. The world of typesetting, image recording, storage and output has been radically transformed in a 25-year period. What's more, the advance of technology as applied to publishing shows no sign of abating.

We can safely predict further improvement in computer to plate transfer methods and in digital integration of text and graphics. We can expect, through lower cost and more versatile electronics, even greater decentralization of the publishing process. The number of networked systems and the opportunity to share resources will increase.

Direct laser imaging of lithographic plates and gravure drums is operational in alpha locations. Interactive color work stations....though prior to Hell's announcement here... I thought not yet economical....are opening new opportunities for creative artists. Tools to capture words and position illustrations are moving from the printing plant to the author's desk.

Systems for library retrieval from relational data bases are being created at several commercial publishers. It is clear from our user survey that electronic publishing has introduced significant benefits in the timeliness and quality of their documentation; further, that having capacity and access invites demand for even greater capabilities and even more access.

It is our view however that "desk top" publishing is the latest desperate attempt to find a stable market for stand-alone PCs, and, in our mind, PostScript is little more than an etch-a-sketch toy for adults.

Despite that assessment, there is little doubt that the proliferation of personal work stations and the expanded computer literacy created by these marketing efforts will help make possible PROFESSIONAL publishing systems producing higher quality, more useful and more timely publications at lower total cost. The real world....our world....will continue to benefit.