Acrobat PDF as

Catalyst for

Convergence

By Bill Zoellick

The following article is reprinted from the weekly analysis of the document software systems market published by CAP Ventures. The weekly analyses are published for subscribers to CAPV's Document Software Strategies Service, which provides market research, analysis, and strategic consulting to companies buying and selling document software systems.

For more information about the service or about the following article, contact Bill Zoellick, Director. His email address is *Bill_Zoellick@capv.com*.

This week Adobe made a number of announcements regarding their new Capture product, which provides new, much more powerful capabilities for converting paper documents to PDF. Capture turns the paper into something more "alive," portable, intelligent, and manageable.

My own association with PDF and the Acrobat product goes back a long ways, to even before it was called "Acrobat," while it was being discussed under the old code name of "Carousel." At that time I was a Senior VP leading the technical side of a company called "Avalanche" (now a hockey team — I couldn't handle the transition). We were working with Adobe to automatically add structure to PDF through the use of SGML. Shortly after that we sold Avalanche to Interleaf, and I was suddenly looking at Acrobat from the perspective of a company that had a competing product called "Worldview." Over the last year, since joining the research and analysis side of this business, I have been watching PDF through the lens of user studies and market research. What I have seen is interesting and worth sharing.

Growth over the Past 3 Years

Acrobat and PDF have not had an easy ride since they first emerged in 1992. In fact, looking at our document management research in early 1996, it looked like PDF was losing ground. The number of survey respondents interested in using it as a preferred, "standard" format was still creeping upward, but as I looked at the level of interest in PDF across different industries, it actually appeared that the focus was narrowing to publishing operations. Early last year, outside of publishing, companies that were looking at viewing solutions were giving strong consideration to HTML. It appeared that HTML could actually begin displacing PDF.

What a difference of year's experience with HTML makes! Every year, in our document management survey, we ask respondents to tell us what format standards they are using or are planning to use for electronic document delivery. We allow users to provide multiple responses if they are using more than a single standard for formatting. Table 1 shows how the responses have changed over the past three years.

Table 1. Delivery Format Standards for Document Management Apps

Delivery Format Standard	1995	1996	1997
None	5%	9%	7%
SGML	36%	24%	26%
HTML	—	33%	49%
PDF	10%	17%	33%
Other	27%	16%	22%

The survey population for these studies consists of companies that either have a document management system in place or that intend to put one in place in the coming year. We're not merely talking about "web sites" here — but about companies that have made a full commitment to a document system. Consequently, SGML usage is relatively high. Not surprisingly, HTML continues to become more important over time. But it is PDF usage that is suddenly growing with particular strength.

We see the same indication of new importance for PDF and Acrobat when we ask these document management customers about the specific client products they are using to view their documents. Acrobat and web browsers are the leading viewing products. As Table 2 shows, once again, 1996 was something of a "hiatus" year for growth in this area, apparently as companies took the measure of the web, intranets, and HTML. The data for 1997 certainly suggest that the market has made its decisions and is ready to start growing again.

website. We also asked them to tell us how the break-down between formats would change over the coming year. The results are in Table 3.

Table 3.

Storage Formats for Web Sites

Document Storage Format	Currently	Next Year
HTML Pages	66%	61%
Word Processing	26%	27%
Relational Databases	21%	21%
SGML	10%	11%
PDF (Acrobat)	7%	12%
Timed Media (voice, video, animation)	7%	8%
Java applets	6%	11%
Other	21%	20%

Not surprisingly, a lot of what is on a website is HTML. What is much more interesting is:

• The percentage of HTML stored on a typical site will probably tend downwards, and almost certainly will not grow

• The amount of Java on a site and the amount of PDF on a site will nearly double

The numbers don't tell us WHY this is happening, but it is pretty easy to guess. HTML has limited expressive power; it does not give people the ability to control the presentation of a document at a very satisfying level of detail. Both Java and PDF offer ways to increase the expressive power of a website. In early 1996 users didn't know this; by now they do.

We get some indirect corroboration of this notion when we turn to look at printing on websites. Our survey finds that, in general, companies are investing in more local and departmental level printing facilities as a result of their intranet investments: people print from their intranets. When we ask about the importance of a matchup between a printed rendition and what a user sees on the screen, the responses are shown in Table 4.

Table 2.

Use of Viewing Products in Document Management Applications

Delivery Product	1995	1996	1997
Web Browser	24%	24%	37%
Adobe Acrobat	11%	11%	29%

What is Behind the Growth?

Since it is reasonable to assume that it was the web and HTML that stalled PDF last year, it is therefore probably useful to look at web and intranet usage patterns for clues as to why PDF is suddenly growing in importance again. In our intranet research, which looks more at website issues than at full, document system concerns, we asked respondents to tell us about the formats that they have stored on their

Table 4.

Match Between Printed Document and Screen Rendition

Must Documents Printed from the Intranet Match Appearance on Screen?	Percent
Yes, always	50%
Yes, sometimes	38%
No, never	11%
Not sure	1%

The respondents clearly attach great importance to the fidelity between the screen and the page. It is reasonable to assume that simply having pages that look like a screen dump of an HTML presentation is not what most of these respondents have in mind; the generally accepted standard for paper presentation is much higher than that. PDF is a way to solve this problem.

The Capture Announcement

As these survey results show, it looks like PDF is finally taking hold as an important expression of "electronic paper." Adobe's persistence in investing in and developing this market appears to finally be paying off. Consequently, the new Capture 2.0 product comes along at an auspicious time.

The obvious benefit of Capture 2.0 is that it has the potential for greatly reducing the cost of converting paper documents into high quality PDF documents. It is set up for use in high speed, production environments, as Adobe's announcement with Cornerstone demonstrates.

Another interesting feature of Capture 2.0 is that the pricing model gives Adobe a way to tie a revenue stream to USAGE of the product. This "pay as you go" model is very important because it allows Adobe to keep the cost of initial purchase under $1000, without forcing them to sacrifice revenues when the product is used in high production operations such as service bureaus. Response to this will be interesting to watch; Adobe could be setting an important precedent that could apply to other document system applications. It is the kind of pricing model that might be attractive to Intranet Service Providers and other vendors set up to do pricing by volume and usage.

Finally, and perhaps most importantly, the introduction of Capture 2.0 coincides with a powerful convergence of image management and document management. It used to be that these were separate markets, serving separate needs. But the distinction between imaging applications and document management applications is now becoming increasingly blurry. We have intranets to thank for this. Once customers can see how disparate systems can be tied together; they are very reasonably asking why their imaging system should be separate from their intranet and from their document repositories.

But truly bringing imaging and document management together can turn out to be harder than one might think, in part because methods of access are so different: management of revisable documents depends in part on full text searching, but document image searching relies on keywords. Acrobat Capture and PDF offer a bridge between these two worlds, enabling scanning and capture of high fidelity images while at the same time opening the captured documents to full text search and other more "intelligent" document operations. Adobe is therefore in a position to accelerate the convergence of imaging and document management, while benefiting from the collision. Adobe's role as catalyst in this convergence is one that other vendors should pay attention to. And it's clear that the vendor attention is there, judging from the list of more than twenty companies that have signed on as part of the Capture 2.0 announcement.