PDF Days Aim to Unlock Power of Format

Despite the fact that PDF viewers are installed on almost every computer—desktop and mobile, and despite the fact that it has been more than five years since **Adobe** gave up control of the PDF specification and the format was adopted as an ISO standard—Duff Johnson thinks the popular document format (no pun intended) is still underutilized.

Johnson, who has spent the majority of his career championing PDF technology, was recently named Executive Director of the **PDF Association**—an organization for which he had previously volunteered. As part of his new role, Johnson helped organize a pair of upcoming North American PDF Days aimed at educating PDF users, as well as ISVs and integrators, on the full potential of the file format. 'These events are different from anything we've ever done before in North America," said Johnson. "Since the PDF Association's inception eight years ago, we have been very focused on promoting adoption of PDF technology primarily among the developer and ISV communities. Now we are taking the value proposition of PDF and presenting it to a wider community. Even though people are using PDFs every day, we don't think they pay enough attention to the technology."

Johnson said members of the ECM community, including ISVs, are more than welcome to attend the PDF Days. They are being in Washington, DC, on Wednesday, Dec. 10, and in New York City on Thursday, Dec. 11. Each event has more than 15 scheduled speakers and starts at 8 AM, running the length of a work day.

The Washington event will focus on government topics and features a speaker from **NARA** (the U.S. National Archives and Records Association) who will discuss his organization's recently issued revised guidance for federal agencies when handling documents. The New York event will focus on the legal and financial markets and feature a panel discuss on "the best possible electronic document format" that will include Harvey Spencer of **Harvey Spencer Associates**, Richard Medina, co-founder of **Doculabs**, and Mark Gross, president and CEO of **Data Conversion Laboratory**.

Cost to attend is $75 for D.C. and $100 for New York. Johnson is expecting 100-150 attendees at each venue, with 150 being the limit. "We had something we called a PDF Day in Seattle last year, but that was basically an adjunct to a technical event and we did it mainly because we had a free location," said Johnson. "We had no sponsors. The upcoming events are really a follow-on to a very successful PDF Day we held recently in Europe.

Duff Johnson, Executive Director, PDF Association

"For the US events, we have 15 sponsors and can accommodate more. Vendors are able to speak in a session at the end of the day called 'Four Minutes with a PDF Vendor,' but all the other sessions are designed to be educational and non-promotional. Even most of our vendor speakers aren't going to be sales and marketing people—they are going to be PDF gurus."

ECM Industry Remiss on PDF

We asked Johnson to specify some on the PDF functions that he sees as being underutilized. "First off, PDF has a digital signature model that is completely open and there is not a better model that I am aware of in the US. PDF also has standardized meta data and annotation capabilities that aren't utilized by most ECM vendors. PDF/A for archiving, as well as PDF/E for engineering drawings aren't used enough. And, most people don't realize you can

embed multiple file types in a PDF and use it like a ZIP."

We had caught up with Johnson previously this year at the AIIM 2014 show where he was beside himself about ECM vendors' poor support of PDF. "I am shocked by the lack of interest in PDF by the ECM world," he told *DIR*. "ECM vendors tend to think about PDF the same way they think about TIFFs.

"I went to every ECM booth I could find [at AIIM] and asked the exhibitors if their software could ingest PDFs. 'Of course we can,' was their response. In fact, some said 80% of the documents being captured into their systems were PDFs. Then I asked, what do you do if the PDF already has notes on it? And everybody gave me a blank stare."

Johnson's issue is that PDF has a standardized way for handling annotations that is typically superseded by the proprietary annotations functionality included in ECM systems. "So, what if a business has a bunch of PDFs that have been marked up by the legal department and then they want to move those documents to an ECM system for archiving?" asked Johnson. "If those annotations don't show up in the ECM system, is that really serving the customer?

"PDF also has a standardized meta data format called XMP. It's been there for years. If ECM vendors placed XMP data in the PDF file, in addition to placing it in their own databases, it would make PDF files completely portable. Unfortunately, because of their history with TIFFs, most ECM vendors just treat PDFs like TIFFs—which they are not; they can be so much more."

The other side of the coin

As a follow-up to our AIIM conversation with Johnson, we talked to pair of ISVs that create viewers for ECM applications. In addition to PDFs, these viewers are utilized on a multitude of other file formats and include annotation capabilities. Regarding standardized PDF annotations, the viewer ISVs offered reasons why they are not typically managed within ECM systems.

"One major drawback of storing annotations within the file is that a new file must be created when each new annotation set is created—either replacing the original file or creating a new version that must be tracked within the system," said Rick Scanlan, director, sales engineering for **Accusoft**. [*Johnson's answer to this in a post-pulication comment: "Not so - annotations in PDF may be individually added, removed, changed, etc., without a problem. They could readily be made, for example, to mirror the annotations stored in the EDMS."*]

"Also, keeping annotation data separate from the image file [like typical ECM systems do] allows for a single version of the file to be stored and provides a way to support multiple annotation sets for each user, track and audit annotation creation and deletion, and enables annotations to be transferred to other systems," said Scanlan. [*Johnson: "PDF does all of this, and in a fully ISO-standardized way."*]

Simon Wieczner, president and CEO of **Snowbound Software**, said the feedback he receives from customers is that the capabilities inherent in PDF often do not meet users' ECM requirements. "Our experience has been that customers are not satisfied with the provided PDF capabilities and work with us or other vendors to get additional functionality," he told *DIR*. "For example, our customers have asked for permission levels that can be set based on different users, the ability to choose between annotation layers, for versioning of different layers of meta data and annotations, and the ability to have tracking data for logging and auditing who has modified the file. These are features that ECM systems with a front-end like ours can offer." [*Johnson: "Each of these features is available in PDF. In otther words - they could implement their features using ISO-standardized PDF technology, indeed, integrate it with their own flavorings. PDF allows it all."*]

Scanlan added that being able to manage annotations for a variety of file formats is important in the ECM market. "Most of our customers need to store a variety of content types (Office, CAD, JPEG, etc.), and they need an approach to annotations that is not tied to a specific document type," he said. "They want one set of annotation tools that can be used for any document type."

It seems the ideal solution might be not ignoring PDF annotations and meta data structures, but rather utilizing them as a supplement to ECM systems' proprietary annotation and meta data management. This would potentially provide users with the best of both worlds—their PDF files would maintain their annotation integrity and meta data, but users would also still have the flexibility and power of their ECM systems' annotation and meta data management.

Of course, this is the stuff Johnson would love to discuss at the upcoming PDF Days. "I think there is a tremendous opportunity for an ECM vendor that decides to ingest PDFs properly," Johnson told *DIR*. "If you treat PDFs like TIFFs, you are missing out on tons of capabilities that PDF has to offer. And if 80% of some users' files are being stored as PDFs today, I think it is time that ECM ISVs started realizing the potential of PDFs."

After the AIIM show, *DIR* also caught up with Thomas Schneck, president of ECM ISV **Docuware**, and we asked him about his company's PDF management strategy. "We are aware of the capabilities of PDF, and we are considering them even more in the future, compared to what we have done so far," he told *DIR*. "But, to really do that, depends mainly on the use cases, the need in the market, and the need by our customers.

"We already have some situations where we are enhancing customers' PDF documents with meta data. This includes support of ZUGFeRD—a German (and possibly future EU) PDF standard for electronic invoices. We support generating as well as reading the required PDF XML metadata. We also support the Spanish Signature standard, both the signature itself and also certain meta data defined by the standard."

For more information: http://bit.ly/PDFDays