

# **Annotating PDFs in Web-Based ECM Systems...Without Altering the Original PDF**

---

**Rick Scanlan, Director, Sales Engineering  
Pegasus Imaging Corporation**

How do you annotate a web-based PDF without altering the original file? It's easy, with the right technology.

Adding annotations to electronic documents is nothing new. It's been done in the electronic content management world for many years. Early adopters of document imaging technology realized quickly that the ability to view and print archived scanned documents is great, but adding workflow with the ability to add notes to documents provided immediate business benefits.

A basic concept of workflow is document review, whether it is a contract that is moving through legal review, an invoice that has been submitted for payment, or an insurance claim that contains a variety of documents that needs to be processed. The documents are scanned, basic indexing information is captured, and the documents are routed to individuals for review, processing and approval. Imaging technology makes this process easy and economical, saving companies significant operations expense.

### **PDF vs. TIFF**

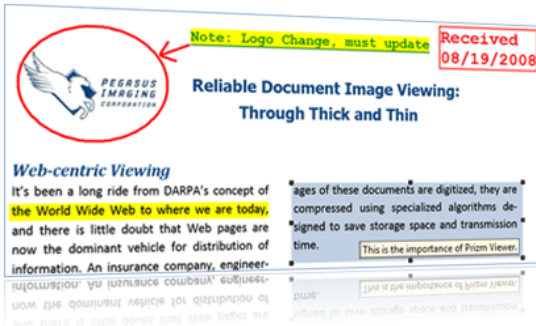
The TIFF file format for document images was the accepted standard for many years. It is still the most widely used format due to its support of excellent lossless Group 4 bitonal image compression and that it is an official public standard. In 100 years you will still be able to read a TIFF file. You likely won't be able to view other image and document file formats that are not official standards.

In recent years, the use of the PDF format has become much more widespread. Adobe certainly made a great business decision to begin giving away the Adobe Reader. The prevalence of the Reader gave the format a huge boost, and most successful scanning applications now support PDF output as an option.

While PDF provides a rich environment of features and is commercially supported, it is not an internationally accepted standard file format. Adobe controls the format, and decides what is added and what is removed, and can change the file format and structure at will. Adobe changes the format a bit with each new Acrobat release in order to provide new features and capabilities. Users need to purchase the latest Adobe technology to take advantage of the new features. And you have to upgrade PDF technology used in other applications to ensure the new format can be used. An official standard format for document archive, PDF/A is available, but it is not yet in wide use. Most PDF documents today are in one of the variety of PDF versions designed by Adobe.

## PDF Annotations for Web-Based Content

The challenge of using PDF in a workflow environment relates to annotations. An annotation is something added to a document. It may be a sticky note, a highlight, or objects such as arrows, circles or boxes.



Annotations are commonly used to attach a note, highlight an area of concern, redact sensitive or confidential information, or make document change requests.

Annotations are frequently used in application such as engineering drawing review, legal research and invoice processing.

Traditional client/server environments provided annotation capabilities. Client software provided all the imaging capabilities the users needed, including the ability to add

annotations to images.

Then the Internet arrived... workers could now work remotely. Widespread access to central document repositories became possible. A web browser became the universal application, with content and application accessed through web servers.

Browser environments are different than client/server environments. While browser technology has come a long way since the early releases of Netscape and Internet Explorer, they still lack certain features that users need for day to day business operations.

Document workflow and collaboration is a great application for web-based applications. Workers in multiple, widespread physical locations can perform document review and processing from any location, even while traveling. Documents can be scanned in central locations and loaded into a company-wide content management system. Remote capture technology can be used in remote office locations, automatically uploading documents to a shared location. Authorized internal users can easily retrieve documents from any location. Contractors and customers can access documents related to their contracts or accounts. Up to date technical documentation can be easily retrieved to perform maintenance and repair operations. Seems like a perfect scenario, until you factor in various image file formats and the need for workflow with document review.

Browsers are great. They display lots of content natively, such as HTML text and JPEG images. However, most documents images are TIFF or PDF, which browsers

don't support by themselves. Technology needs to be added to browsers to support these additional formats. A variety of TIFF viewers are available, some free, some licensed. Most companies choose commercially licensed viewers that they know will be supported and upgraded.

Adobe® Reader® is the software most often used to view PDF files in a browser. The Reader is free and is easily downloaded and installed. The Reader is designed for individual use, with a basic viewing feature set. The Adobe® Acrobat® family of products provides additional features, including the ability to create PDFs and add annotations to existing PDF files. The free Reader allows you to view existing annotations but does not let you add them.

While Adobe Acrobat products let you add annotations, the software is designed to work with "local" files. The file may have been downloaded from a web site, received in email, or created locally. Acrobat by default saves annotations within the file itself. If the file was downloaded from a web-based content management system the new file must then be uploaded and stored in the system. Two files are now stored, the original and the annotated version. The original document, especially in the case of scanned documents, typically needs to be permanently stored as the legal representation of the original document. If additional users need to review and annotate the document, several versions of the document might be stored, increasing storage requirements and creating database bloat. Accurate document versioning can also be a problem, as well as managing user annotation rights.

### **A Solution... Store Annotations Separately**

Annotating PDF files using an overlay approach, where the annotations are stored separately from the original document, is a much better approach than editing the document file. The annotations are still associated with the document, but the original file is not touched and multiple stored copies of the document are not needed.

Shared annotations, stored separate from the document, allow multiple people to view, add, delete and edit the annotation set without altering the original document file. Individual annotation sets can be stored for each user, with document management permissions controlling which annotations a user can view and edit, and whether they can add their own annotations to the review process. A single copy of the original document can be stored. Annotation content can be managed and searched.

Prizm® Viewer, from Pegasus Imaging Corporation, is an example of a browser-based application that is ideal for working with web-based content management systems. Prizm Viewer installs in the local users browser, similar to the way Adobe Reader is installed. Prizm Viewer adds considerable viewing and annotation capabilities for a wide variety of image and PDF file types. One of Prizm's greatest strengths is its ability to annotate PDF files retrieved from content management systems, without altering the original file. When annotations are added to a document, only the annotation content is transmitted to the web server to be stored. Prizm Viewer annotations can be stored directly in any content management system, indexed along with the PDF document. When a user retrieves a PDF from the system, Prizm Viewer is automatically called to display the document. Prizm quickly checks the server for annotations permitted for the individual users, and automatically displays the annotation content with the document.

Prizm Viewer has been used with a wide variety of content management systems in many different organizations, including companies accessing financial, engineering, and legal documents in industries including aerospace and heavy equipment, financial and insurance, and government applications such as emergency management coordination, patent research and approval, and others. Prizm's additional capabilities such as Image List Files to bring together separate but related documents for viewing and printing, and its programmable API make Prizm Viewer the choice of many large organizations. With over 1,000,000 installations across more than a thousand organizations, Prizm Viewer provides the viewing and annotations capabilities that organizations need to access web-based content management systems.

### Summary

Easy access to an organization's documents is critical for business success. Web technology can enable widely spread employees access to documents. The ability to add annotations to documents through workflow and content management systems continues to be an important tool, regardless if documents are accessed through traditional client/server systems or through web-based applications. The special challenge of not only viewing PDF files but also sharing document annotations without modifying the original document requires careful consideration and planning. Organizations should strongly consider technology that provides the best of both capabilities.