Effective Personal-Identity and Metadata Redaction Techniques for Subsequent E-Filing

When you e-file a PDF document, you may be providing more information in that document than you can see via your PDF reader software. Some redaction techniques used when e-filing are ineffective, in that the text intended to be hidden or deleted can be read via a variety of techniques. And, because information about the document, called "metadata", is also stored inside the document, it is often viewable as well. Examples of metadata and hidden data include the name and type of file, the name of the author, the location of the file on your file server, the full-sized version of a cropped picture, and prior revisions of the text.

E-filers must use extra care to make sure that the PDF documents to be submitted to ECF are fully and completely free of any hidden data which may contain redacted information. The protection of sensitive data can be compromised if improper redaction techniques are used. Here are a couple of examples of sensitive-data visibility issues:

- Highlighting text in black or using a black box over the data in MS Word or Adobe Acrobat will not protect the data from being able to be seen.
  See: http://www.pdfforlawyers.com/2005/05/pdf_redaction_s.html

- Previous revisions and deleted text may be able to be seen by manipulating an Adobe Acrobat file. See:
  http://www.computerworld.com.au/index.php/id;1097572794;fp;4194304;fpid;1
  http://blog.didierstevens.com/2008/05/07/solving-a-little-pdf-puzzle/

Fortunately, there are effective means of eliminating this metadata from electronic documents. Probably the simplest method is to omit the information from the original document and save the redacted version with a new name. For example, a Social Security number can be included as XXX-XX-1234.

The court does not endorse any specific method, and the responsibility for redacting personal identifiers rests solely with the parties.