Making Digital Facsimiles of Documents
Part 1: Overview

This article is an introduction to the process of making digital facsimiles of documents in a form that is suitable for publication on the Web.

The term document as used here refers primarily to material that is printed on paper, although many of the same considerations apply to photographs, microfilm, xerographic copies, as well as other kinds of material, such as fabric samples.

The Process

There are four main steps in making a digital facsimile of a document:

• Scanning the pages of the document to produce corresponding images.

• Touching up the images.

• Placing the images on pages with a page layout program to produce a digital document

• Producing a Portable Document Format® file (PDF) from this digital document.

The PDF [1] is the digital facsimile.

Note: Adobe Capture® can scan directly into a PDF, but it lacks the flexibility to produce good results for all but the simplest of documents.

The steps above can be performed by different individuals at different locations. For example, one person may produce the scanned images and transfer them to another person who does the page layout and produces the PDF.

Each of the steps above may involve a variety of processes and decisions, depending on the document being copied. These will be covered in subsequent articles.

Hardware Requirements

• A Windows or Macintosh computer with sufficient hard disk space to store image files — at least 1 GB of free space. Speed is an important consideration because some processes involved in producing digital facsimiles are computationally intensive.

• A flatbed scanner. Almost any scanner of current vintage has the capabilities needed to produce digital facsimiles. The least expensive scanners, however, may be slow and physically flimsy.

Software Requirements

• An image-manipulation program that can handle the "plug-in" that comes with the scanner. Such programs are bundled with most scanners. The ones that come with the least expensive scanners, however, may be slow and limited in capability. Photoshop® or Photoshop Elements® is recommended.

• A page-layout program that allows images to be placed on document pages and manipulated easily. Adobe PageMaker® is recommended.

• Adobe Distiller®, which is needed to produce PDFs from page-layout output.

Image-manipulation programs are part of the usual software repertoire and are bundled with scanners. Page-layout programs are less common. Distiller, which is part of the commercial Adobe Acrobat® package, usually is a separate acquisition, although it is packaged with some page-layout programs.

Note that only a scanner and the software that comes with it is needed to produce images. The production of PDFs requires more.

Reference

1. Griswold, Ralph E. *Portable Document Format*, 2001:
   http://www.cs.arizona.edu/patterns/weaving/pdf.html

Ralph E. Griswold
Department of Computer Science
The University of Arizona
Tucson, Arizona

January 8, 2003; last modified April 1, 2004