

Serving PDFs on the Web

Leonard Rosenthol
Chief Technology Officer
PDF Sages, Inc.



Copyright©2001-2003, PDF Sages, Inc.

Static Files



Copyright©2001-2003, PDF Sages, Inc.

Overview

- Static Files
- Forms
 - Simple Data Exchange
 - Server-Side Processing
 - Form Merging
 - Stamping/Overlaying
 - Dynamic Production
- Document Conversion
- Document Assembly
- Protecting/Securing your work



Copyright©2001-2003, PDF Sages, Inc.

Static Files

- If you just have some documents that you want to make available to your users, you can simply upload them to your server and link to them like any other type of file.



Copyright©2001-2003, PDF Sages, Inc.

Improving the static

■ Linearization aka “Fast Web View”

- A process of saving or preparing a PDF file in such a way that it can be viewed dynamically (on demand - “a page at a time”) in a browser rather than having to have the whole PDF downloaded before anything can be seen.

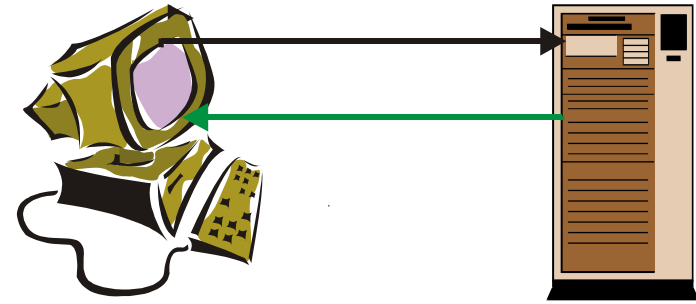
■ ByteServing

- A feature of HTTP/1.1 and modern web servers that provides the “back end” for linearization to work.
 - Current versions of Apache, IIS, WebStar all support this

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Simple Data Exchange aka. Client-Side Merging

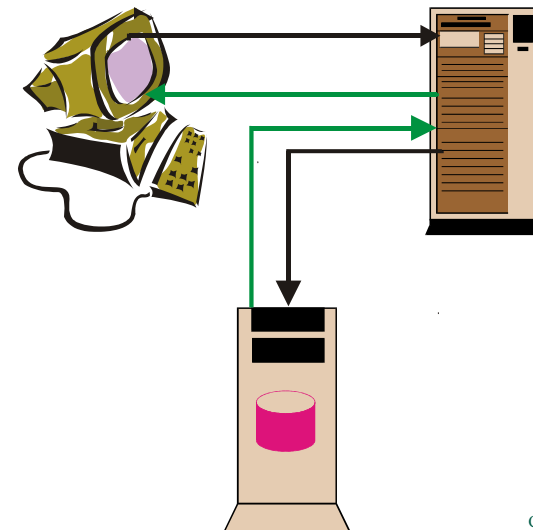


pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Forms

Something missing?



pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

A bit more in depth...

- Acrobat submits HTML, FDF, or XFDF via HTTP
 - Optionally you could use an HTML form & submit in std. HTML Format
- CGI/ASP on server receives the data
- It then processes it
 - potentially interacting with a database
- Possibly returns new data
 - in either FDF or XFDF format
 - in the same or a different PDF document



Copyright©2001-2003, PDF Sages, Inc.

HTML vs. FDF vs . XFDF

- HTML form posting is a standard
 - Supported by all CGI and ASP tools.
 - It does not, however, provide support for all Acrobat features.
- FDF is the “old standard” for transferring PDF form data.
 - Requires the use of the Adobe FDFToolkit to parse the files
 - generating them is easy enough with any good text processor
- XFDF is an XML-based version of FDF.
 - This means you get the benefits of FDF, with the ability to use your favorite XML tools.
 - Limited with Acrobat 5, but Acrobat 6 offers MUCH enhanced XFDF support



Copyright©2001-2003, PDF Sages, Inc.

What do I need to make it work?

- Web Browser
- Acrobat 3.5 or later
 - Acrobat is the only PDF viewer with forms support, which were introduced in 3.5
- Web server w/CGI or ASP support
 - Pick your favorite OS and server software along with integration language (Perl, VB, Python, PHP, etc.)
- Optionally you may have a database or other server-based tools to manipulate/massage the data



Copyright©2001-2003, PDF Sages, Inc.

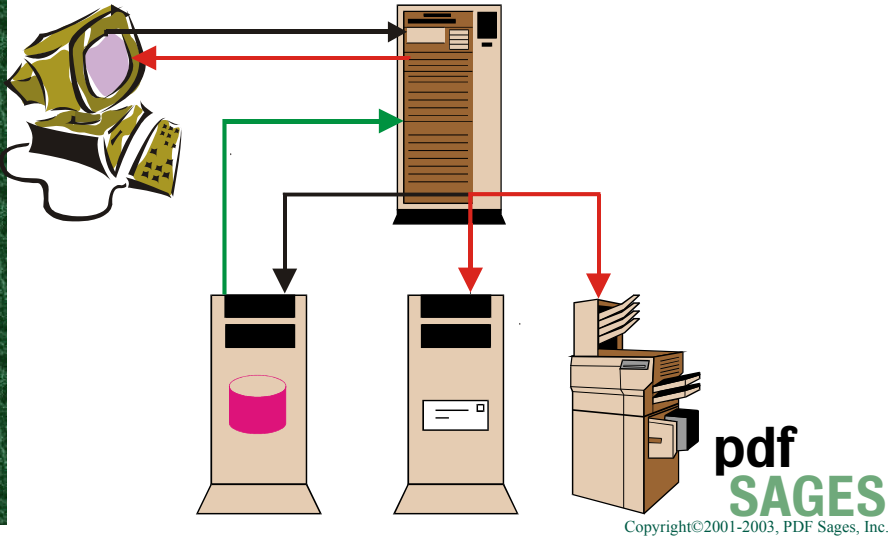
Pros and Cons of Simple Exchange

- Pros
 - Doesn't require any 3rd party tools
 - Cheap
 - No need for admin access on server
 - Support for Acrobat 6 features today
 - Rich text in fields!
- Cons
 - Requires Acrobat
 - Mac OS & Windows
 - No Unix, Palm, PocketPC, etc.
 - Reader uses can NOT save the filled in form
 - Browser incompatibility issues
 - Can't return data into separate window or frame



Copyright©2001-2003, PDF Sages, Inc.

Server-Side Processing



Server-Side Options

- Form Merging
- Stamping/Overlaying
- Dynamic Production

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Pros and Cons of Server-Side

- Pros
 - Does NOT require Acrobat
 - So you can support Unix, Palm, PocketPC, etc.
 - Reader uses CAN save the filled in form
 - NO browser incompatibility issues
 - CAN return data into separate window or frame
- Cons
 - Requires 3rd party tools
 - \$\$ is involved
 - How much is determined by your specific needs
 - » OS platform, integration language, database support, etc.
 - NEED admin access on server to install/config

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Form Merging

- Start with a standard PDF form
- Use a 3rd party tool to fill in the form fields
 - From FDF
 - From XML/XFDF
 - Direct from a Database
- Optionally “flatten” the fields
- Return the filled-in form to the user

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Available Tools

- Adobe Document Server - Adobe
- FDFMerge – Appligent
- PDFToolkit – ActivePDF
- iQuickSEDPDF - SEDTech
- iText



Copyright©2001-2003, PDF Sages, Inc.

Stamping/Overlaying

- Start with any PDF – no form fields required!
- Use 3rd party tools to “stamp” data onto PDF
 - May require creating new document and copying original first, then overlaying.
 - NOTE: these products will remove all interactive elements (bookmarks, hyperlinks, etc.) from the source PDF’s
 - Some of these tools use XML-based description languages to define where to place the data.
- Return new PDF to the user



Copyright©2001-2003, PDF Sages, Inc.

Pros & Cons of Form Merging

- Pros
 - Easy to setup and understand
 - No need to be a programmer
 - Some tools can automate the entire process
 - Not just text, but also checkboxes, buttons, etc.
 - Flattened forms can be linearized
 - NEVER linearize a PDF form!
- Cons
 - Limited formatting options for data
 - For non-flattened forms that will be compatible with Acrobat 4, you are limited to the Base14 fonts
 - Limited layout options
 - No reflow, extra white space, borders, etc.
 - Limited non-Roman text support



Copyright©2001-2003, PDF Sages, Inc.

Available Tools

- Adobe Document Server - Adobe
- StampBatch – Appligent
- PDFToolkit – ActivePDF
- PDFlib – PDFlib, GMBH
- PDFever – Perl Studio
- ReportLab – ReportLab
- iText



Copyright©2001-2003, PDF Sages, Inc.

Pros & Cons of Stamping

- Pros
 - Greater flexibility of formatting & layout options
 - May support non-Roman text
- Cons
 - Requires manual determination (trial & error?) of where elements are placed on the page.
 - PDFlib has a nice “Blocks” feature to solve this
 - Little to no support for non-textual elements
 - May require some programming

Available Tools

- Adobe Document Server - Adobe
- ReportLab – ReportLab
- FOP – Apache Project
- XEP – RenderX
- iText
- PdfTeX

Dynamic Production

- Maintain layout and “template” content in a database or other content repository
 - eg. XML, MIF
- Combine the “template” with the data
 - eg. XSL-T
- Process the new content in conjunction with styling information
 - eg. XSL:FO
- Return the new PDF to the user

Pros & Cons of Dynamic Production

- Pros
 - Greatest flexibility of formatting & layout
 - Pagination and document reflow
 - May include tabular formatting, charts, etc.
 - May support non-Roman text
- Cons
 - LOTS more setup/configuration work
 - Will require programming

Document Conversion



pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Available Tools

- ActivePDF DocConverter - ActivePDF
- AdLib eXpress - AdLib Systems

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

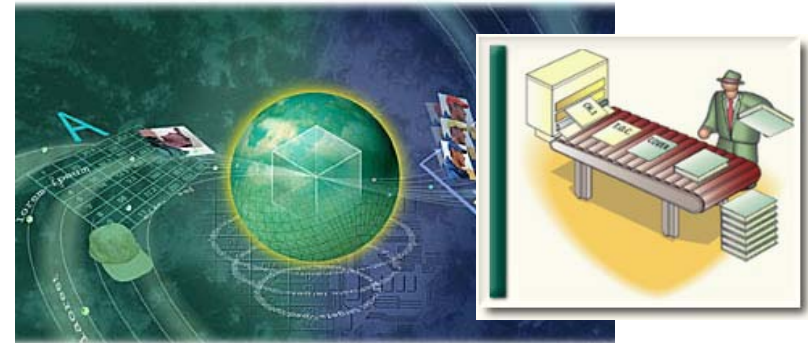
Document Conversion

- Start with existing Office or other documents and convert them to PDF on the fly to deliver to the user
 - Allows you to make sure you have the latest versions of the documents
 - Doesn't require your document authors to have Acrobat or other tools
 - Allows you to potentially use tools for mail-merge or similar features

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Document Assembly



pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Document Assembly

- Create new PDF documents from existing PDF documents
 - either static or dynamically prepared
 - Appending/merging
 - Extraction/Splitting
 - Extra material
 - Inclusion of custom data (stamping)
 - Table of Contents (TOC)
 - Bookmarks

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Protecting & Securing your PDFs



pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Available Tools

- Adobe Document Server – Adobe
- AppendPDF/AppendPro – Appligent
- PDFToolkit – ActivePDF
- PDFlib – PDFlib, GMBH
- iText

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Protecting & Securing PDFs

- Whether you start with a “premade” PDF, create dynamic documents or convert from a Word file, you may want to make sure that your content is protected
- Digital Rights Management (DRM)
 - Don’t Allow Copying
 - Don’t Allow Printing
 - Full Access Control
 - Timeouts & Expiration
 - Pay for use or purchasing
- One thing you can NOT do is prevent the file being downloaded. Like all net content, it has to be downloaded before viewing.

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.

Acrobat Security vs. 3rd party

- Acrobat
 - Included with every copy of Acrobat (and Reader), so no extras need be downloaded
 - Not very secure
 - Limited free functionality (WebBuy is extra)
- 3rd party
 - Requires user to install a plugin
 - More secure
 - More features to choose from

Reader Extensions

- Add functionality to Reader on a PDF-by-PDF basis
 - Commenting
 - form filling & saving
 - etc.
- Requires Adobe Document Server + Reader Extensions
- Priced for Intranets or large installations - \$75K

Available Tools

- WebBuy - Adobe
- FileOpen (Personal) Publisher - FileOpen
- PageVault – Authentica

Wrap Up

- Now is a GREAT time to putting PDFs online, as there are millions of users worldwide(!) that can view your documents.
- Start with static documents - it's free and easy!
- But begin thinking about how you can improve your users experience with your documents by adding server-side options.

Q & A

Q & A

pdf
SAGES

Copyright©2001-2003, PDF Sages, Inc.