



**Color Feature Archive Procedure**

**For**

**PDF Archive/Electronic  
Tearsheets**

**Prepared by  
Jeff Dodd**

**Revision 3.01  
May 9, 2005**

## Revisions

---

REVISION	RELEASE DATE	COMMENT
1.00	October 22, 2004	Initial release.
2.00	December 10, 2004	After Lancaster started to deliver PDFs from production.
3.00	Apr 29, 2005	Integration of new Excel macro and use of PDF Enhancer.
3.01	May 9, 2005	Inclusion of Inside TV

---

## Table of Contents

1. Introduction .....	4
2. PDF Page Creation.....	4
3. File Assembly.....	5
4. Feature Layout/Editorial Map.....	6
5. File Retrieval .....	7
6. Editing.....	7
Appendix A — Transferring Files .....	13
Appendix B — File Assembly .....	14
Appendix C — Flat File Creation .....	16
Appendix D — Scripts and Macros Used.....	17

## 1. Introduction

---

Until the contents of the editorial map(s) can be integrated into the scripts assembling the color feature pages, the pages need manual intervention to be assembled for the PDF archive. This must be done on a timely basis as a version is delivered to advertisers through the electronic tearsheets system. This document describes the steps, scripts and deliverables for creation of the weekly files.

The process has been scripted to minimize the switching between systems. It presumes the user:

- has access to UNIX servers MAG3 and SGINFS;
- has Acrobat Professional;
- has the ISI Bookmark toolkit installed as a plugin to Adobe Acrobat; and
- has the Excel macro to assist with the formatting of the bookmark CSV file.

The main sections of the document describe the workflow for producing the archive and the appendices provide the details of the scripts involved should there be a failure in the process or the list of products needs to be changed.

## 2. PDF Page Creation

---

This section is informational and is provided to detail the workflow. No action should be required of the user for this part of the process. Program details are provided in **Appendix A —Transferring Files**.

The color production house (R.R. Donnelley in Lancaster) delivers PDF files of all pages (advertising and editorial) and covers to a shared production directory (**Exchange**) on SGINFS. The **Digest** files are normally delivered early Monday morning (around 2:00am). **Inside TV** files are normally delivered on Tuesday mornings. Big Book files are normally delivered on Wednesday mornings.

Once an hour, a script, `/usr/tvg/support/jdodd/MoveArchivePages.sh` is executed (running under the `edadmin` user `crontab`) that watches the delivery directory `/usr5/exchange/To_Radnor/to_Archive`. The script looks for files found ending with “.PDF.” If the files are delivered to a sub-directory, the files will not be found.

To assure that the files are named to meet production needs, the script renames the cover files to sort correctly for batch processing. For example, the “C300.PDF” file is the third version of cover one, not cover three, “C3.PDF”. If there are multiple first covers, they may be named C300AC.PDF and C300BC.PDF. These files may need to be renamed manually. UNIX meta characters, such as the ampersand (“&”) and space, are also removed from the file names. Other meta characters may appear under manual circumstances and may need to be renamed by hand.

Stable files are sent to `/usr/work` on MAG3. After each file is copied from SGINFS, it is renamed to “**moved.(filename)**.” The files are saved should there be file corruption in the build process and the originals are needed.

At the end of the **cron** session, e-mail is delivered to a definable group of users, notifying them of the file transfer. An example is shown below.

```

From: Generic Editorial Administration login [edadmin@tvguide.com]      Sent: Tue 10/5/2004 4:18 PM
To: Patricia Muller; Marta Rubel; Peggy Nolan; Jeff Dodd
Cc:
Subject: PDF files moved to archive workarea at 04:18:25 PM

=====
38 PDF files found at Tue Oct  5 16:15:03 EDT 2004 ...
Renaming TD42C300.pdf to TD42C1003.pdf
Renaming TD42C100.pdf to TD42C1001.pdf
Renaming TD42C200.pdf to TD42C1002.pdf
  TD42F0001.pdf sent at Tue Oct  5 16:15:09 EDT 2004
  TD42F0003.pdf sent at Tue Oct  5 16:15:14 EDT 2004
  TD42F0008.pdf sent at Tue Oct  5 16:15:19 EDT 2004
  TD42F0010.pdf sent at Tue Oct  5 16:15:25 EDT 2004
  TD42F0011.pdf sent at Tue Oct  5 16:15:30 EDT 2004
  TD42F0012.pdf sent at Tue Oct  5 16:15:35 EDT 2004
  TD42F0013.pdf sent at Tue Oct  5 16:15:41 EDT 2004
  TD42F0014.pdf sent at Tue Oct  5 16:15:46 EDT 2004
  TD42F0016.pdf sent at Tue Oct  5 16:15:52 EDT 2004

```

If the file is not successfully delivered (because the remote partition is full or other file transfer problem), the program will note the files to be left for the next session. The script will exit and reattempt transfer one hour later.

### 3. File Assembly

Again, this section is informational and is provided to detail the workflow. No action should be required of the user for this part of the process. Program details are provided in **Appendix B — File Assembly**.

Once every two hours, a script, `/netserv/jdodd/AutoFeature.sh`, is executed (running under the `edadmin` user **crontab**) that watches the contents of the work directory `/docs/work`. The script looks for files starting with “T” and ending with “.PDF” and the product character in the second position, “D” for Digest, “B” for the large format weekly, “S” for Special Interest Publications and “I” for **Inside TV**.

The script counts the number of pages from the same product. The file name with the largest page number is examined, such as “TD12F0048,” and compared to the number of pages present. If there are fewer pages than that, the script continues to the next product. The script also assesses the available space in the work area. If less than half of the space is available, the script will exit. (It assumes that all pages will be duplicated in the combined file.)

A batch file is built of the pages in sequential order. Two applications are used as part of this procedure, Appligent **AppendPDF Pro** and Apago’s **PDF Enhancer**. Both applications are server versions and are compiled for AIX.

When it is estimated that all of the pages are present, the script starts to assemble the archive. The individual page files are created for print reproduction and have “trim” and “bleed” areas displayed. This will be trimmed from the displayed page. The production files demand large storage requirements, two-to-three megabytes per page for Digest and ten-to-twenty megabytes for Inside TV. The first task is to “reduce” the PDF file with PDF Enhancer. Through the down sampling of images and other conversions, each page is reduced to 80-95% of the original size. If the page compression fails, the original page remains and will be assembled into the archive.

A bookmark is assigned to each page as it assembled into the main document. If the page name is found in an {ISSUE}-{PRODUCT}.list, i.e., **051505-D-list**, the bookmark is extracted and inserted into the PDF assembly batch file.

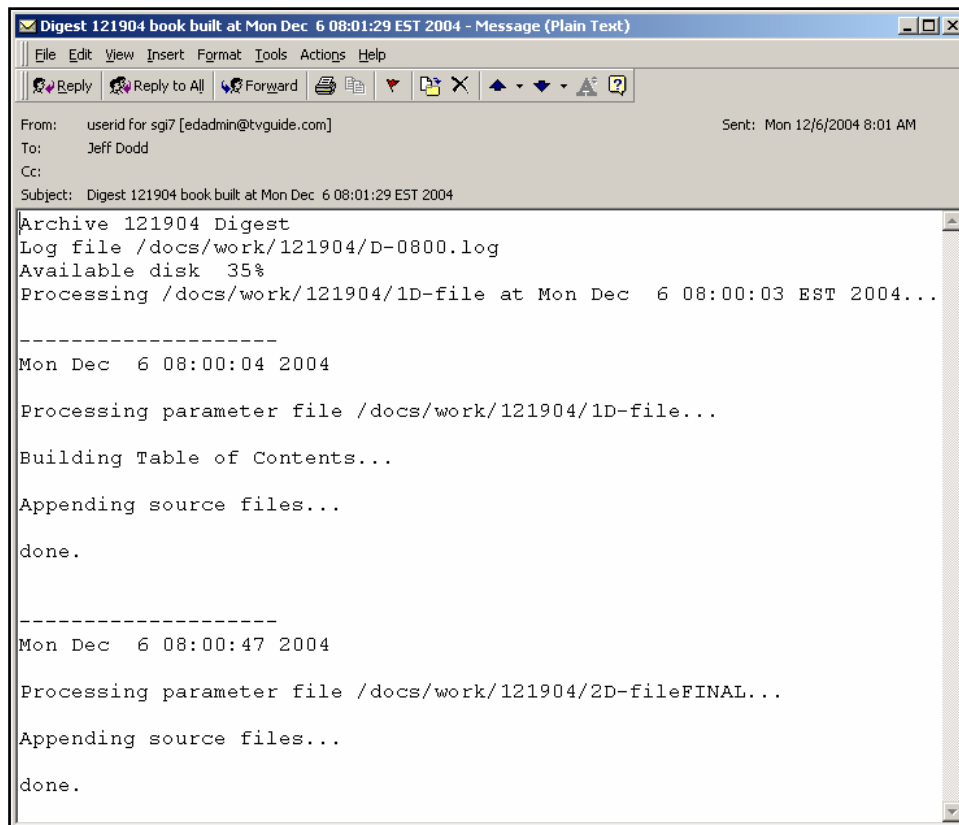
If a bookmark is not found, the script constructs a bookmark consisting of the page (folio) number (from the file name) and the “logical” page number (the number of the page relative to the start of the document).

The script attempts to maintain the correct left-to-right page orientation after the covers are inserted and if regional color pages are found. Cover 2 should be on an even-logical page number to appear in the spread adjacent to page 1 since the spread is frequently an ad spread. Blank pages have been created for each product. **BlankDig.pdf** is sized to be the trimmed size of a digest page. **BBBlank.pdf** is sized to work with the large format weekly and **Inside TV**.

The script confirms an issue data directory is present under the path of the PDF archive web space, **/docs/archive/pcspdf**. This area is published as <http://magazinearchive.tvguide.com>. The merged, finished PDF file is moved to the issue directory.

To prevent this script from running after the archive file has been built, the script first checks for the presence of the first AppendPro batch file (**1{PRODUCT}-file**). If one is present, the script exits.

A message is sent upon completion of the build, notifying the user that the magazine has been assembled.



```
✉ Digest 121904 book built at Mon Dec 6 08:01:29 EST 2004 - Message (Plain Text)
File Edit View Insert Format Tools Actions Help
Reply Reply to All Forward Print Copy Paste Undo Redo
From: userid for sgi7 [edadmin@tvguide.com] Sent: Mon 12/6/2004 8:01 AM
To: Jeff Dodd
Cc:
Subject: Digest 121904 book built at Mon Dec 6 08:01:29 EST 2004

Archive 121904 Digest
Log file /docs/work/121904/D-0800.log
Available disk 35%
Processing /docs/work/121904/1D-file at Mon Dec 6 08:00:03 EST 2004...

-----
Mon Dec 6 08:00:04 2004

Processing parameter file /docs/work/121904/1D-file...

Building Table of Contents...

Appending source files...

done.

-----
Mon Dec 6 08:00:47 2004

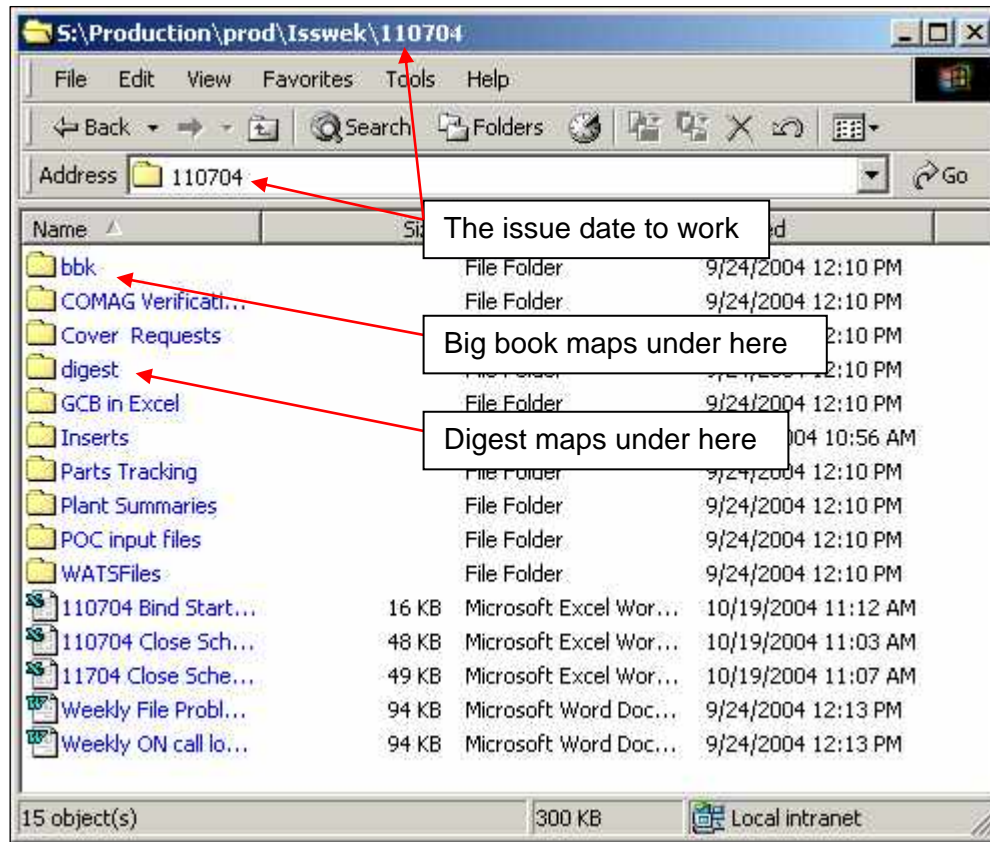
Processing parameter file /docs/work/121904/2D-fileFINAL...

Appending source files...

done.
```

## 4. Feature Layout/Editorial Map

The color feature section layout or editorial maps (for Digest and Big Book) are completed in Microsoft Excel. The maps are located on the NT servers. The “issue week” directory is under (rad1fp2)S:\Production\prod. The issues that are being worked are displayed under that directory. Double-click on the issue that is to be built. This will display several production files for the issue.



Under each *product* directory will be a **map** directory. Inside of that directory will be all of the versions of the map. (“Final” in the name does not necessarily mean the last version; look for the newest file.) Select the latest one and copy it to a directory on the local hard drive. (This just makes accessing the file easier.)

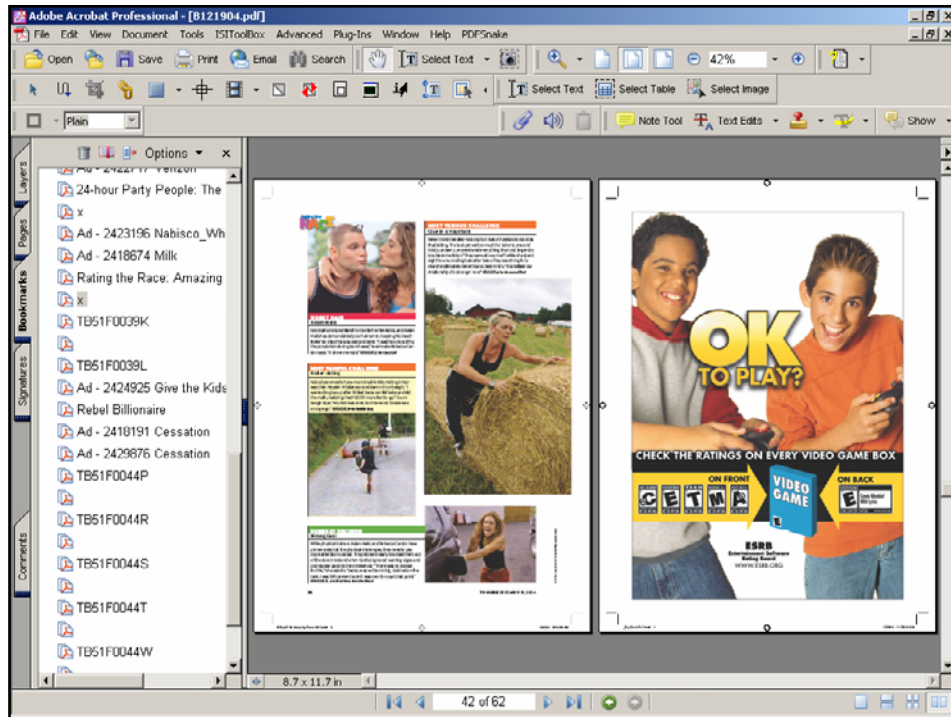
## 5. File Retrieval

To modify then reapply the bookmarks, the file must be pulled to the desktop. This can be done in two ways: by opening the file via the web browser (which will, under most circumstances, open Acrobat in a browser window) and save the file to the local disk; and by running an FTP session to MAG3 to get the file.

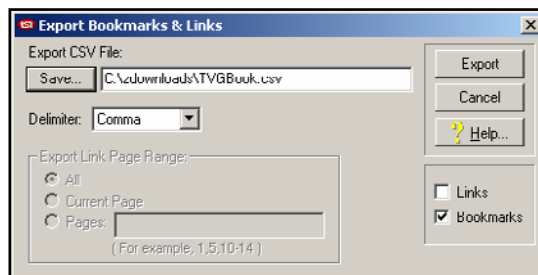
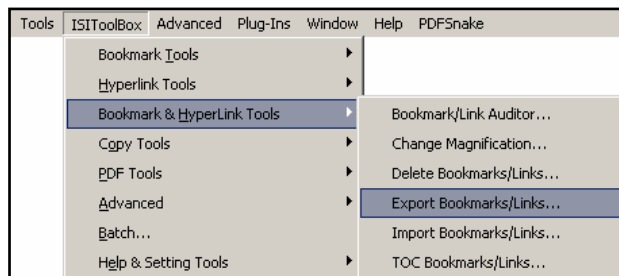
## 6. Editing

To edit the bookmarks, the user will need Adobe Acrobat Standard or Professional with the ISI Toolkit (version 5.1 or greater) plugins installed. Open the PDF file. The file should display the first page, to the right of the screen, with the “raw” bookmarks displayed at the left.

- By default, the bookmarks are all black. In the “raw” file, you may see blank bookmarks, a few with “x’s” and the bookmarks generated from the batch file.



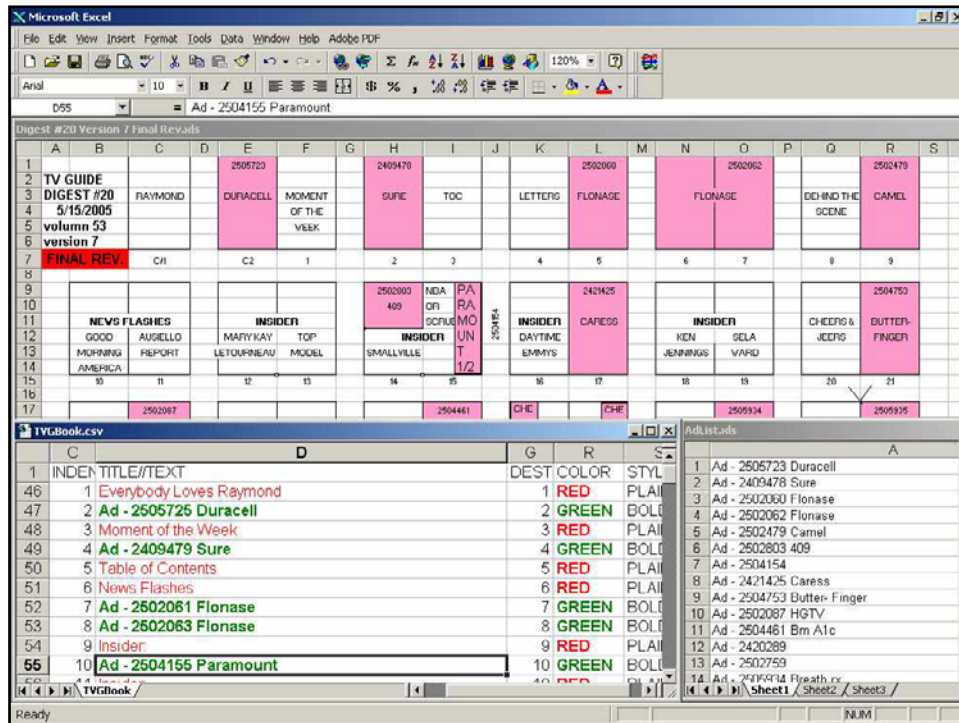
- If you wish, you can scroll through the pages and edit the bookmarks for split pages by inserting an “f,” or “l,” for the position of the ad.<sup>1</sup>
- Export the bookmarks using the ISIToolbox plugins. Use the standard name<sup>2</sup> for the file of **TVGBook .csv**.



<sup>1</sup> A “shorthand” is available to prepare to apply two bookmarks to the same page. If the bookmark is preceded by an “f,” (“first”), the macro that is run later in MS Excel will copy the current line for page Ids and such, put the ad IO on the first line and “Partial Edit” on the second bookmark. If the bookmark is preceded by an “l,” (“last”), the macro will copy the current line and put the ad IO on the second line and “Partial Edit” on the first bookmark.

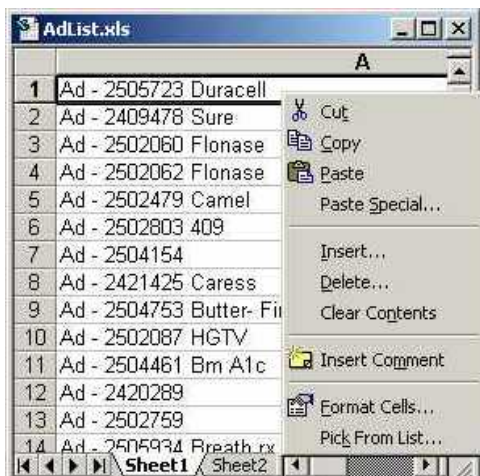
<sup>2</sup> The file name for the CSV file is arbitrary, but the macro, used in later steps, needs to look for a consistent name. Likewise, the path where the file is written is arbitrary, in the example “C:\Zdownloads.” This can be reassigned, but the macro, used in later steps, needs to look for a consistent name.

- Load the map for that product in MS Excel. Run the **PrepBookmarkList** macro. This will first search the map, row-by-row, for 7-digit numbers starting with a “2.” It will then search down the column below it searching for the ad description. This will create a new table **adlist.xls**.
- The macro will then open the **TVGBook.csv** file and arrange the map, ad list and **.csv** sheets on the screen. (This presumes the **.csv** file is stored in the path illustrated above.) The macro will process the each line in the map, deleting blank bookmarks and those marked with the short-hand “x” (selected to be deleted). Duplicate bookmarks for split pages are created as directed by the first character on the line. Finally the standard feature bookmarks (“Rough Review,” “Cheers & Jeers” and table of contents) and list of ads are inserted at the top of the description column and the rows are hidden. This allows for “fast lookup” of often repeated bookmarks. Other columns not needed by the user are also hidden.

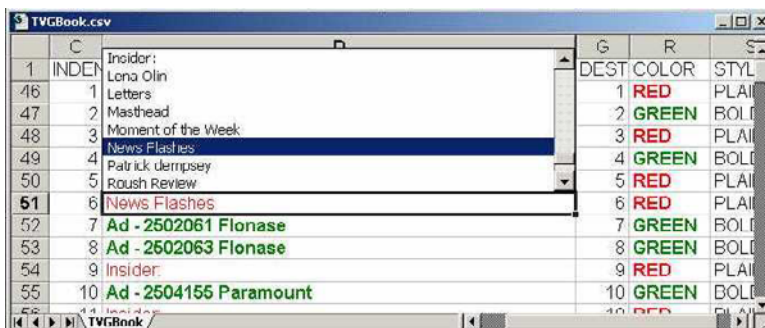
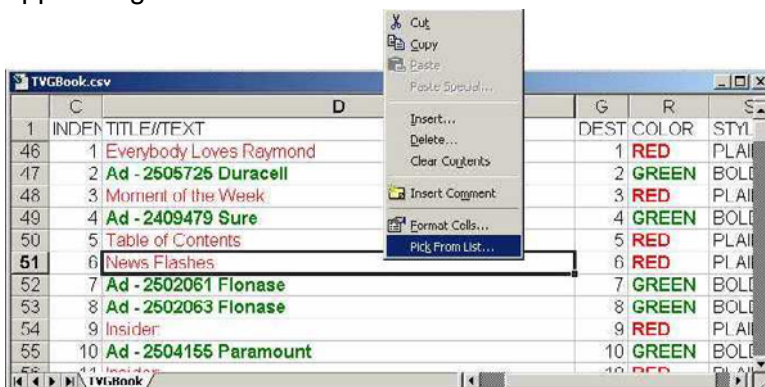


- Embedded into the CSV file will be formulas to change the colors of the bookmarks based on the title. If the first four character of the string starts with “Ad<space>-“, the parameter for the column is set to GREEN (for the bookmark color). In the next column when those conditions exist, the weight of the type parameter is set to BOLD. Otherwise the bookmarks are RED and PLAIN for editorial content. Conditional formatting is applied to the parameter columns and the title column to mimic the results in the MS Excel display.
- The description of the bookmark is entered in the title column. This can be done three ways:
  - Typing the description into the cell (for editorial feature titles);

- Selecting a single cell from the **adlist.xls** table with the IO number and advertiser name, copying the cell to the clipboard (with a CNTL-C on the keyboard, a right click of the mouse on the cell and a menu will appear, via the icon on the toolbar or “Edit,” “Copy” from the menu bar) and pasting the content into the appropriate cell of the **TVGBook.csv** file;

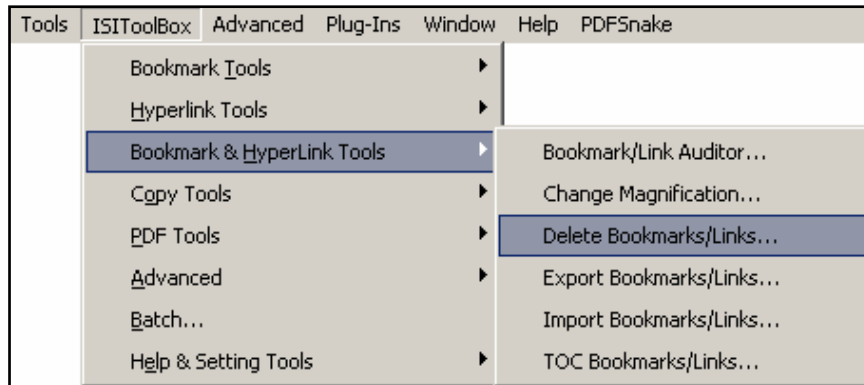


- Right-clicking on the destination cell and selecting “Pick from List...” This displays an alphabetically sorted list of all values displayed in the title column. Since the standard feature names and entries from the ad list are hidden at the top of the column, all will appear together.

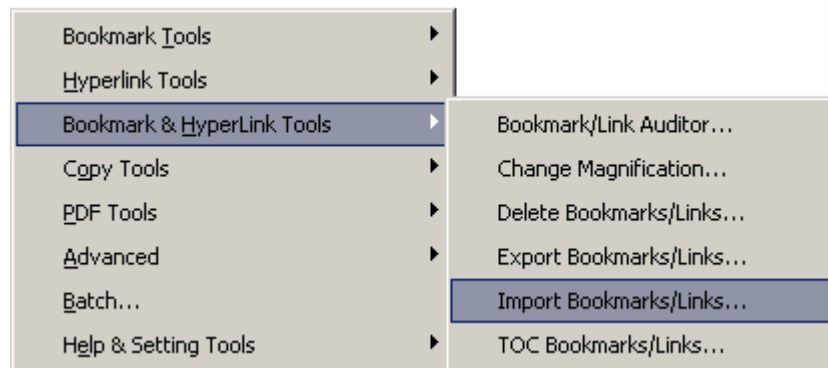


- Complete the entries for all columns. If two cells are needed for split pages, execute the “split page” macro. This will copy all of the entries (some of which are hidden) into the new line.
- If there is a continuation of the story on a following page, the row can be deleted. Select the entire row by clicking on the row number in the far left of the window. From the menu at the top of the screen, select “Edit” then “Delete.”
- When finished save the file. Make sure it is still a CSV file.

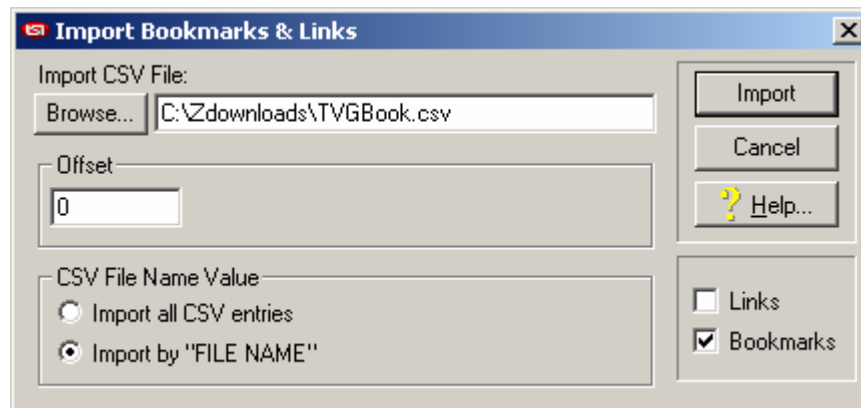
- Return to Acrobat. Before the new bookmarks can be imported, the existing ones must be deleted, otherwise there will be duplicates embedded in the file. Using the ISIToolbox plugins, delete all of the bookmarks.



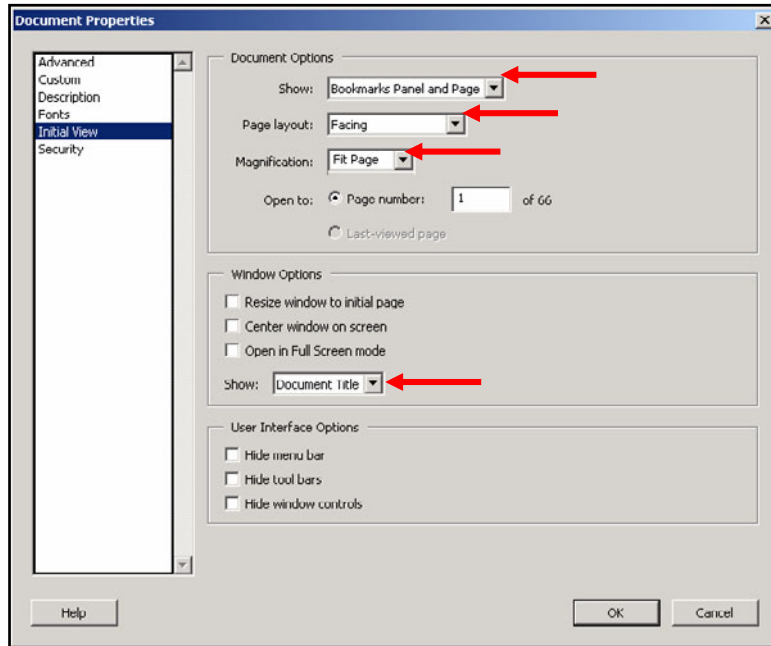
- Import the bookmarks from the modified **.csv** file using the ISIToolbox plugins.



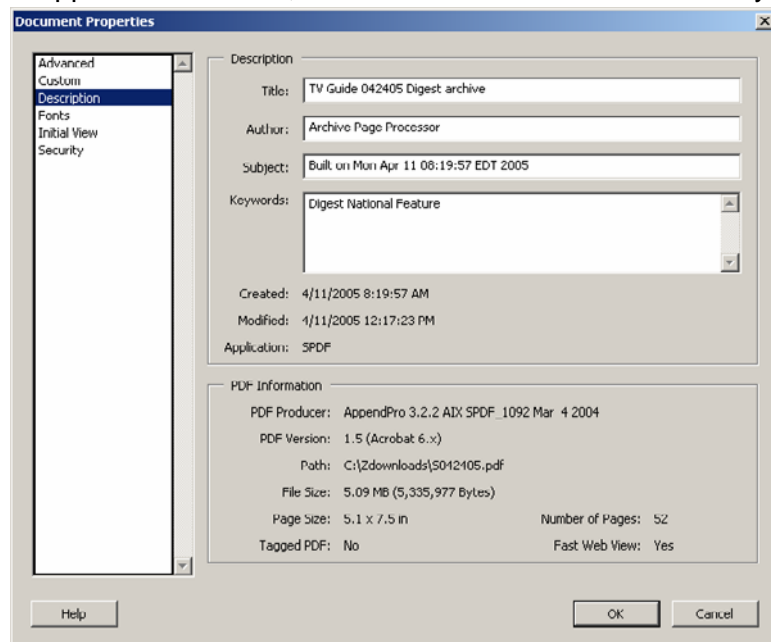
- Using the same file name makes the process very simple.



- Set the default Document properties for the title and file display. Click “File, Document Properties” and when the window appears, click the “Initial view” tab. Set “Show:” to “Bookmarks Panel and Page,” “Page layout” to “Facing” and “Magnification:” to “Fit Page.”



- Under “Window Options,” set “Show:” to “Document Title.”
- Check the Document “Description.”
- If AppendPro has run, these fields will be set automatically, as shown below.



## Appendix A — Transferring Files

---

- The script that moves the files is running from `/usr/tvg/support/jdodd/` and is named `MoveArchivePages.sh`. The version, as of May 1, 2005, is 1.03.
- It contains a routine for “exception” name processing for covers and the **Inside TV** pages. The pages typically arrive with the following naming convention:
  - TD18C102.PDF for Digest pages
  - TB18F025G.PDF for Big Book pages
  - NT18B0088.PDF for Inside TV pages
- The first two characters represent the product. To support script compatibility, the **Inside TV** product string “NT” is renamed to be “TI.”
- The third and fourth character at the issue week.
- The fifth character represents the type of page. “C” represents covers and “F” represents feature pages. To support script compatibility, the **Inside TV** page string “B” is renamed to be “F.”
- The next one-to-four characters represent the page number. In the case of pages, the four digits represent the folio number of the page. In the case of covers, a single digit is the number of the cover (2, 3 and 4). Three digits represent the cover variation: news stand, subscription and hotel covers. When there are numerous first covers, such as Football or Simpsons, the two digits following the “1” may increment the count.
- Letters following the numbers represent regional variations. “25A,” “25B” and “25C” are three versions of page 25 running in distinct markets. There have not been more than 26 regional pages in a single edition.
- Since the script could start running in the middle of file delivery, the script lists files with the oldest first, calculates the checksum of a file, sleeps for 5 seconds, then recalculates the checksum. If the values match, it is assumed the file is stable and can be processed.
- Known meta-characters are stripped out by a sed command. This has not been a problem of late.
- There is a script in `/usr/tvg/support/jdodd/` named `Unmove.sh`. This is used to remove the “moved” prefix from each file name to allow them to be re-transmitted to MAG3.

## Appendix B — File Assembly

---

- The script that assembles the files is `/netserve/jdodd/` and is named **AutoFeature.sh**. The version, as of May 1, 2005, is 2.01.
- When **AutoFeature.sh** runs, it builds the book in smaller 40-60 page chunks, with a table of contents page at the beginning of each chunk. This is required to allow bookmarks to appear on each page. The TOC page is cut off when the smaller parts are assembled into a single large document. The smaller chunks are required, as AppendPro will have a memory fault if it attempts to combine more than roughly 70 pages at one time.
- The bookmark list of {ISSUE}-{PRODUCT}.list can be created with a script in `/netserve/jdodd/` named **MakeMarker.sh**. This will create a flat ASCII file with page numbers, a pound sign (#) and the page description to be inserted into the bookmark.
- The script checks the available disk space before starting to work. The assumption is that two complete copies of the pages will exist in the workspace during the assembly of the file. If there is less the 50% of the disk space available, the program may go into “cheat mode.” Here it will put temporary files under the web archive disk area. Upon completion, the temporary area is deleted.
- Documentation for Appligent **AppendPDF Pro** and Apago’s **PDF Enhancer** are available as PDF files in `S:\IT\Application_Reference_Docs`, in **AppendProGuide.pdf** and **PDFEnhancerGuide.pdf** respectively.
- When the PDF pages are enhanced, the images are compressed to 300 dpi JPEG files, with high-quality compression and conversion to a Ycc-RGB color space. This provides reasonable quality for later printing with significant compression.
- The pages have the bleed areas cropped according to the type of page. The margins on the digest front cover are different that the inside covers, which are different than the pages. The instances of multi-fold pages have been infrequent enough that no configuration has been established.
  - **Digest cover 1:** left margin=1 inch, bottom margin=1 inch, width=5.1 inches, height=7.5 inches
  - **Digest cover 2:** left margin=0.4 inch, bottom margin=0.3 inch, width=5.1 inches, height=7.5 inches
  - **Digest cover 3:** left margin=1 inch, bottom margin=0.3 inch, width=5.1 inches, height=7.5 inches
  - **Digest cover 4:** left margin=0.625 inch, bottom margin=0.25 inch, width=5.5 inches, height=7.375 inches (includes the spine)
  - **Digest pages:** left margin=0.5 inch, bottom margin=0.5 inch, width=5.2 inches, height=7.5 inches (all pages are the same, on left and right pages)
  - **Big Book pages and covers:** left margin=0.5 inch, bottom margin=0.5 inch, width=7.7 inches, height=10.75 inches. (Separate files are created to allow the “calculated” file names to be determined.)
- PDF Enhancer processes each page via an XML configuration file, identified as a “.PEC” file. The files can be coded by hand, however it is easier to create the file via the MS Windows GUI and export the XML. This can be transferred to the AIX host for use. The Windows version can also be used to prototype/test configurations before delivery to AIX.

- To create or modify a PEC file, click “Targets.”

## Appendix C — Flat File Creation

While it is not essential to create the flat file in advance, it is convenient to have as many of the bookmarks in place when the file is initially created. The file can be created after the feature close and before the page proofs are delivered. A script has been created to assist with the file generation.

- Load the MS Excel map.
- Log into **MAG3** and **cd ~jdodd**. (I use Hummingbird X-Terminal)
- Launch **MakeMarker.sh** with an argument of the product, “D” or “B.” By default the map is created for the issue two weeks out. If you want to do next weeks issue, enter the argument of “1.” For example, **MakeMarker.sh B 1** will guide you through the creation of a map for next week.

- Enter the bookmark for the page at the prompt. You can select one of the standing items from the list. As new items are added for this issue, they are added to the list of items to select from.
- As ad IOs are entered, the advertiser name is added to the list. This is helpful in replication the description when multi-page ads are booked in the feature section.
- When there are partial page ads, mark the page as an ad. The editorial part of the page can be marked later.
- When there are regional pages (multiple versions of a page), enter “2” for “No Label.”
- The script will exit when something has been entered for all pages.

## Appendix D — Scripts and Macros Used

---

Below are the scripts used, as of this writing.

- When the high-res file is complete, FTP it to the web space on MAG3.

The “S” file is then moved to the **sginfs:/Horsham/xyvision/e-tearsheets/{issue}** transfer directory. A schedule task on MS SQL Server moves the files to the web server in Tulsa. Make sure the permissions are 777 on the directory and the file.