

Creating a PDF file containing multiple outputs

Spencer Renyard, Phastar, London, UK

ABSTRACT

One of the required components of Phastar's programming environment is the ability to combine table, figure and listing (TFL) files into a single PDF file. The business requirements were collated and reviewed, including reusability and formatting expectations. Potential solutions involving SAS and external applications assessed for suitability against the requirements. Of the solutions investigated only one met the key requirements of reusability without manual intervention, along with the ability to include a hyperlinked Table of Contents (ToC) and hyperlinked bookmarks containing the first title of each TFL. SAS and templates are used to create the individual TFLs as PDF files. TFL metadata is used to control the order of the TFLs in the combined PDF files, with the first title taken from the metadata and used to populate the ToC and bookmarks. The user controls the creation of the combined PDF file by calling a SAS macro.

INTRODUCTION

As part of the internal programming environment one of the required components identified was the ability to create a single PDF file containing the individual table, figure and listing (TFL) files produced by the SAS® programs. This paper collates and reviews the business and client requirements, the collection and review of information for potential solutions and the analysis of these potential solutions, the selection of one solution, and its implementation. The solution presented has been implemented in a Windows Server 2012 environment using SAS 9.4.

REQUIREMENTS

A review of the business and client requirements was performed. This provided the framework when assessing potential solutions.

- Easy to use, e.g. simple SAS macro call
- Repeatable – is there a line command option?
- Scalable / Performance
- Are there cost or licensing requirements?
- Can a subset of outputs be produced?
- Interleave files containing tables and figures
- Run on Windows Server 2012
- Hyperlinked Table of Contents (ToC) using output title
- Bookmarks displaying output title
- Respect formatting options, e.g. paper size, orientation, of source files
- Respect page numbering of source files
- Order files using output number – may not be based on filename

RESEARCH

There are many applications and services ranging from £free to £lots which allow you to easily produce a PDF file from an input file. Some even provide the facility to run multiple files in a single action. Far fewer applications allow the user to compile a PDF file from multiple input files and fewer still guarantee that the output will also include bookmarks and ToC or hyperlinked ToC. Vendors' websites for these applications are often light on information which could facilitate an informed decision on whether to try/use/buy the software

POTENTIAL SOLUTIONS

Several options involving SAS and the various ODS destinations were explored, including using the Document destination and replaying ([1]), and creating PostScript, RTF and PDF files. All of these options required additional steps to create the final file, some within SAS and outside of SAS. A number of external tools were evaluated. Of the tools assessed, most lacked the ability to meet one or more of the requirements listed above, with absence of the ToC and a command line interface being the primary shortcomings.

CHOSEN SOLUTION

One of the external tools investigated, PDF Combine by CoolUtils, included the functionality to meet all of the requirements for a PDF creation system. This included the ability to create a hyperlinked ToC and bookmarks based

PhUSE 2016

in TFL title information, merging PDF files whilst respecting the formatting of the source file, and a command line interface. Used in conjunction with SAS-produced PDF files containing the TFL outputs, this was the solution chosen.

IMPLEMENTATION

SAS ODS in combination with standardized template code is used to create PDF files containing the output from each TFL program. The name of the output file, along with the output type, number, titles and footnotes are stored as metadata in a file external to SAS.

The process to create the combined PDF file is controlled using a SAS macro call, thereby simplifying the interface for the user. The macro includes a parameter to specify a subset of the outputs based on the output number, e.g. 11.1 will select all TFLs whose number starts with 11.1. The macro accesses the TFL metadata to ascertain the file order based on output number, and also obtain the text for the first title line for each output for use in the ToC and bookmarks. Once all the information on file order and bookmark has been processed, the macro creates a Windows batch file containing the text to run PDF Combine, enabling the user to create a combined PDF file containing the required TFLs without the need to know how to use the software employed. The process can also be replicated.

CONCLUSION

The chosen implementation allows users to quickly and repeatedly combine individual TFL output files into a single PDF file with a hyperlinked ToC and hyperlinked bookmarks containing the text of the first title for each TFL.

REFERENCES

References go at the end of your paper. This section is not required.

[1] Coar, W. "TFLs: Replaying Rather than Appending", PharmaSUG 2013, Paper PO16

[2] CoolUtils.com <https://www.coolutils.com/PDFCombine>

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Spencer Renyard
Phastar
Unit 2, 2A Bollo Lane
London W4 5LE
Work Phone: 020 77183 7061
Email: spencer.renyard@phastar.com
Web: www.phastar.co.uk

Brand and product names are trademarks of their respective companies.