



SS04: Post-mortem logs in R

PHUSE US CONNECT 2023 – Orlando, FL

Teckla Akinyi

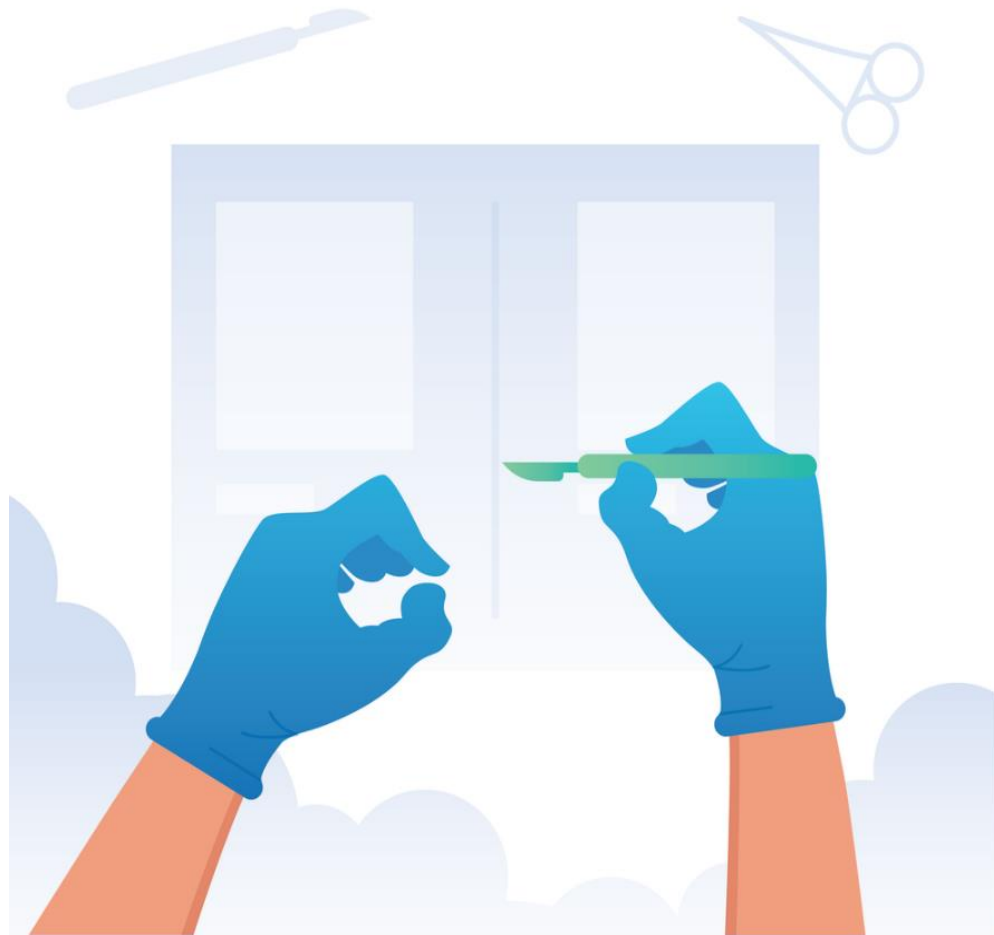
Sr. Programmer

Janssen Pharmaceutical Companies of Johnson & Johnson

The opinions in this slide deck and oral presentation are my own and do not necessarily reflect the views and policies of Janssen Pharmaceutical companies of J&J



Post-mortem logs in R



<https://easyretro.io/blog/post-mortem-versus-retrospectives/>

- ❑ Common process across industry.
- ❑ Analyze log files of scripts that are run for: errors, warnings and other key messages on script execution
- ❑ Practice: extract these messages into a concise navigable file e.g., excel, rtf...company choice

Objective: Share a snippet developed in R to extract messages from log files generated by both SAS and R programming languages into an Excel sheet.

{logrx} versus sas generated log files

SAS: Lines printed in sequential execution order with In-line key words of interest

{logrx}: Organized in sections of interest, keywords as headings or subheadings

```
297
298 data cmpr3(drop=col:);
299     set cmpr2x;
300     CMINDCPR=catx(';',of col:);
    71
ERROR 71-185: The CATX function call does not have enough arguments.

301     run;
```

WARNING: The variable 'col:'n in the DROP, KEEP, or RENAME list has never been referenced.
NOTE: The SAS System stopped processing this step because of errors.
WARNING: The data set WORK.CMPR3 may be incomplete. When this st
NOTE: DATA statement used (Total process time):

real time	0.00 seconds
user cpu time	0.00 seconds
system cpu time	0.00 seconds
memory	889.59k
OS Memory	39328.00k
Timestamp	02/13/2023 04:48:21 PM
Step Count	197
Page Faults	0
Page Reclaims	12
Page Swaps	0
Voluntary Context Switches	12
Involuntary Context Switches	0
Block Input Operations	0
Block Output Operations	272

```
302     proc sort data=cmpr3; by usubjid cmseq; run;
```

Used Package and Functions	
{package:base} data.frame, c, print, cat, environment, library, message, warning, stop	
{package:dplyr} select, arrange	
{package:magrittr} %>%	
{package:rlang} inform	
Program Run Time Information	
Start time: 2023-02-15 10:51:27 EST	
End time: 2023-02-15 10:51:29 EST	
Run time: 2 seconds	
Errors and Warnings	
Errors:	
log error	
Warnings:	
log warning	
log warning2	
Messages, Output, and Result	
Messages:	
log inform	
log inform 2	
log message	
log message 2	
Output:	
[1] "log print"	
[1] "log print 2"	
log catlog cat 2<environment: 0x00000250e0f96ea0>	
Result:	
NULL	

parseLogs(args...)

Age old solutions of macros in SAS reading text log files and exporting a filtered file(rtf, excel etc.)

So why?

Developed in R (version 4.2.2)

Reads both {logrx} and SAS generated files

```
parselogs(logdir = "C:/Users/TAkinyi/Desktop/Yangu2/PHUSEABC",  
          select_file = "*.log$",  
          sas_file_ID = c("SAS Institute Inc| SAS Version"),  
          sas_key = c("ERROR|WARNING"),  
          add_r_sxtn = NULL,  
          Sheet_nm = "sheet name",  
          outnm = "Compound123.xlsx")
```

args(parseLogs) – developed in R version 4.2.2

Argument	Remark	Description
logdir	Required	Path to the input directory, defaults to the working directory
select_file	Optional	For user-selection of log file to be checked only one log at a time. Default to all logs in directory.
sas_file_ID	Required	Identifiers of a sas file, typically a string in header of program that also prints to log text file.
sas_key	Optional if no SAS logs are expected	Key words that user wants to be extracted from SAS files. The list is extensible for user to add other key words per company requirements.
add_r_sxctn	Optional	Used for R files to extract other sections beyond Errors, Warnings & Messages e.g section with masked function. If not specified or left at default then only errors, warnings & messages are extracted.
Sheet_nm	Optional	ID to be used as the excel sheet name can be file name or studyid if looking through different study dirs. Default is "sheet name"
outnm	Required	Name for the excel workbook. Can specify full path or just name. If only the desired file name is specified file will be stored in the path specified in the logdir argument.

Live demo with some example calls



Known Issues/Enhancement Opportunities



<https://www.presentermedia.com/powerpoint-animation>

- ❑ The source code for the function is currently a large function.
- ❑ Currently written to be sourced in script.
- ❑ Addition of a sequential counter to arrange messages in the order of occurrence.
- ❑ Argument constraints:
 - add_r_sxtn* can only be used when parsing an individual {logrx} file.
 - select_file* can only accommodate one file at a time.
 - sas_file_ID* is always required

!!!!This function **has not** been peer tested!!!!

Opportunities are endless in enhancing parseLogs(args)

```
parselogs(logdir = "C:/Users/TAkinyi/Desktop/Yangu2/PHUSEABC",  
          select_file = "*.log$",  
          sas_file_ID = c("SAS Institute Inc| SAS Version"),  
          sas_key = c("ERROR|WARNING"),  
          add_r_sxtn = NULL,  
          Sheet_nm = "sheet name",  
          outnm = "Compound123.xlsx")
```

The source code is available in the following github repository on a read access:

<https://github.com/tkakinyi/phuse2023/tree/main>