

AD07 - Using Python to compile CSR

Speaker: Volodymyr Loskutov, Statistical Programmer
Labcorp, Kyiv, Ukraine

labcorp
Drug Development

+



What is Clinical study report?

- Clinical study report(CSR) is the final step of clinical trials before submitting the results to the regulatory agencies.it provides many details about methods and results of a medical trails. Usually this is a very long document with some tables and figures.

For example: Clinical Study Report (CSR) for AstraZeneca trial 5077US/0049 has 70 tables, 10 figures, 8000 pages



Clinical Study Report (CSR) Template

<https://clinicalstudytemplates.com/>

Template scope and purpose

This Clinical Study Report (CSR) Template is specifically designed to be used in conjunction with the Clinical Study Protocol (CSP) Template.

The structure and sections as well as their order and content follow the International Council for Harmonisation (ICH) E3 guidelines for CSRs, as recommended by the Food and Drug Administration (FDA). As such, this CSR Template is the foundation for an "integrated" full report of any study with a therapeutic, prophylactic, or diagnostic agent (i.e., drug or treatment) conducted in patients or healthy volunteers.

Note that, if conducting a study with healthy volunteers, these may be referred to as "subjects" instead of "patients", except where subjects have become patients due to adverse events (AEs).

Per the ICH guidelines, the clinical and statistical description, presentation, and analyses are to be integrated into a single report (i.e., the CSR), incorporating tables and figures into the main text of the CSR or at the end of the text, with appendices containing such information as the CSP, sample case report forms, and investigator-related information, information related to the test drugs/investigational products (also referred to as study products), including active control/comparators, technical statistical documentation, related publications, patient data, as long as technical statistical details, such as derivations, computations, and listings, are inserted per computer output. Note that the CSR of a study should *not* be created as a separate clinical and statistical report. Although this CSR Template was developed for the most common efficacy and safety clinical studies (e.g., as clinical trials), the basic principles and structure can be adapted for other studies, such as clinical pharmacology studies (e.g., biomarker of exposure [BOE] studies). Depending on the study type, for such studies, a less detailed report might be acceptable.

- For more information on ICH guidelines for efficacy guidelines, visit <https://www.ich.org/page/efficacy-guidelines>

CSR Study Number: XXXXX

Version: X.X

Date: dd-MMM-yyyy

CONFIDENTIAL INFORMATION

Company name here

Page 2 of 103

Problem

- CSR (Clinical study report) may go through multiple rounds of recompiling during the development. Usually this is a manual process
- CSR contains many tables and figures. Each time you need to update them, you must open the rtf document with updated table and replace the old table in CSR. For 1 table, I will spend near 5 minutes
- Sometimes you have only sas7bdat dataset instead rtf documents and you need to wait till programmers will create rtf tables



Initial task

- Create utility to compile CSR.
- Utility may insert tables or figures on special places
- Use sas7bdat datasets to create tables
- Use jpeg images to create figures
- Read titles and footnotes from the file (Mock shell)

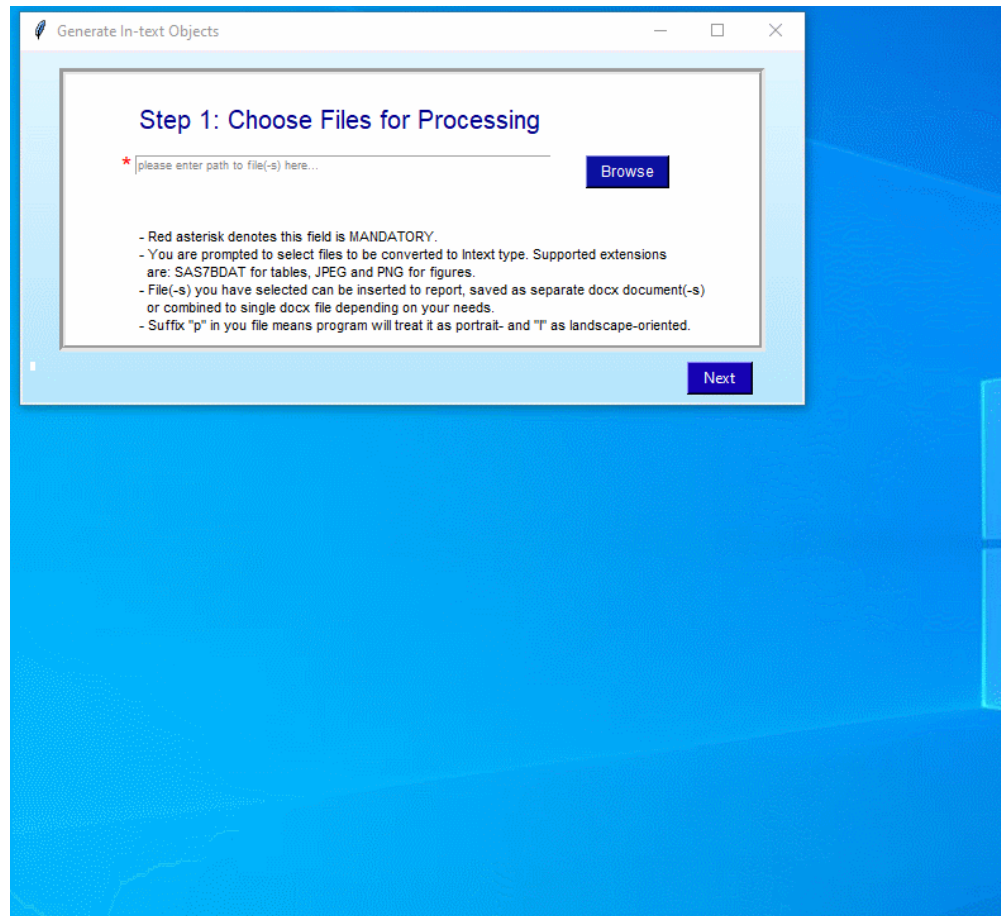


Solution

- Develop the Windows application on Python 3.7

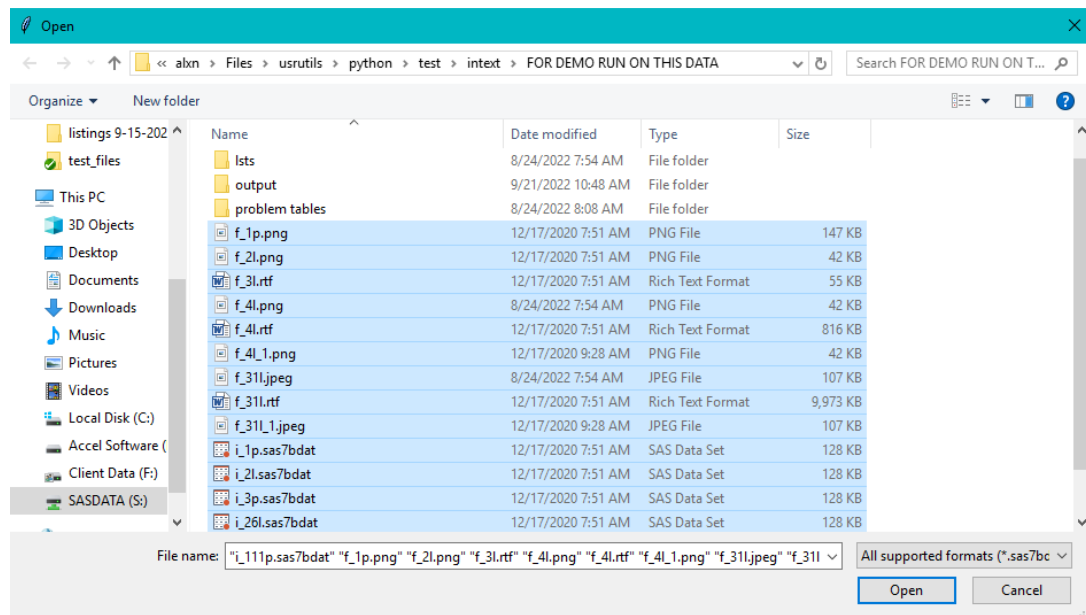
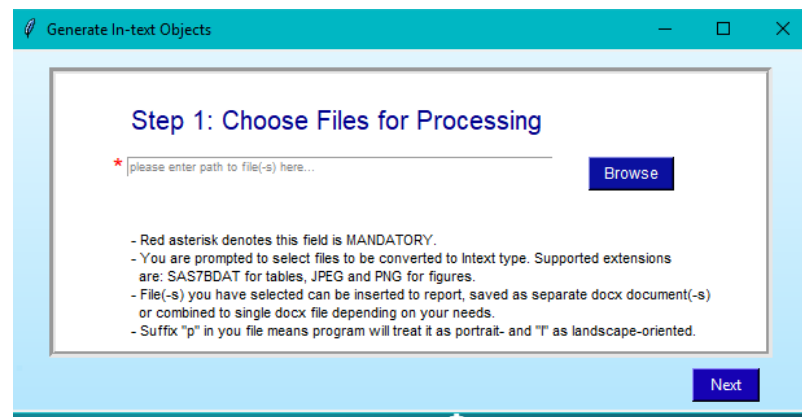
Create simple graphical user interface

Add log window and progress bar



Select files to insert

- multiple files may be chosen
- Sas7bdat datasets used to create tables
- Jpeg and png images used to create figures



Select Mock shell file

Mock shell file has information about:

1. Table numbers
2. Titles
3. Footnotes

Generate In-text Objects

Step 2: Select In-Text mock

* please enter path to file here...

- Red asterisk denotes this field is MANDATORY.
- The Intext mocks you have selected will be used to confirm the presence of the file(-s) from previous step within these mocks.
- Title, population and footnotes from the Intext mocks will be used for creating Intext object(-s) from the file(-s), selected on previous step.

Figure 1
EORTC QLQ-30 Fatigue Subscale: Percentage of Patients with a 10 Point Improvement from Baseline, by Visit (Full Analysis Set)

ADQS.FASFL='Y' and PARAMCD = 'EORTCGL' and ANL01FL = 'Y' and CRIT1FL='Y'.
The plot will be similar to figure 14.2.2.2.1.

Note: EORTC QLQ-C30=European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire-Core 30 Scale.

Figure 2
EORTC QLQ-30 Physical Functioning Subscale: Percentage of Patients with a 10 Point Improvement from Baseline, by Visit (Full Analysis Set)

ADQS.FASFL='Y' and PARAMCD = 'EORTFSL' and ANL01FL = 'Y' and CRIT1FL='Y'.
The plot will be similar to figure 14.2.2.2.1.

Note: EORTC QLQ-C30=European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire-Core 30 Scale.

Figure 3
EORTC QLQ-30 Fatigue Subscale: Percentage of Patients with a 10 Point Improvement from Baseline, by Visit (Full Analysis Set)

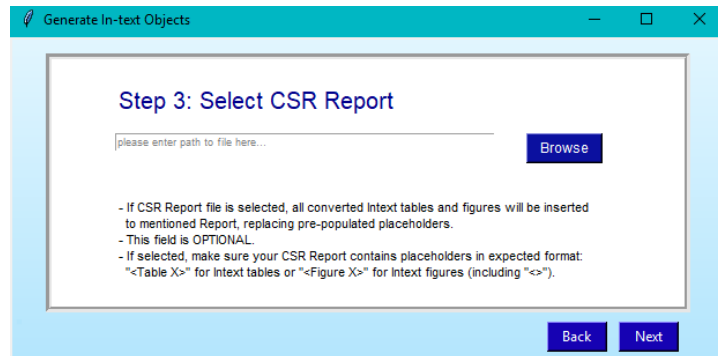
ADQS.FASFL='Y' and PARAMCD = 'EORTSS1' and ANL01FL = 'Y' and CRIT1FL='Y'.
The plot will be similar to figure 14.2.2.2.1.

Note: EORTC QLQ-C30=European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire-Core 30 Scale.

Select CSR

CSR is a text document with special tags (<Table 1>, <Figure 3>).

Tags will be replaced on tables and figures.



A little bit of text appears here. A little bit of text appears here. **A little bit of text appears here.**
A little bit of text appears here.

Some text appears here.

Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here.

<Figure 1>

Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.

<Table 3>

A little bit of text appears here. A little bit of text appears here. **A little bit of text appears here.**
A little bit of text appears here.

Results

IMAGINE THIS IS A REPORT.

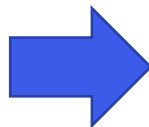
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.

A little bit of text appears here. A little bit of text appears here. **A little bit of text appears here.**
A little bit of text appears here.

Some text appears here.

<Table 1>

Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here.



IMAGINE THIS IS A REPORT.

Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.

A little bit of text appears here. A little bit of text appears here. **A little bit of text appears here.**
A little bit of text appears here.

Some text appears here.



Table 1: Analysis Data Sets (All Randomized Patients)

	Med1 (N=125)	Med2 (N=121)	Total (N=246)
Number of Randomized Patients	125 (100)	121 (100)	246 (100)
Number of Patients in the Full Analysis Set	125 (100)	121 (100)	246 (100)
Number of Patients Excluded from the Full Analysis Set	0 (0.0)	0 (0.0)	0 (0.0)
Number of Patients in the Per Protocol Set	124 (99.2)	120 (99.2)	244 (99.2)
Number of Patients Excluded from the Per Protocol Set	1 (0.8)	1 (0.8)	2 (0.8)
Number of Patients in the Safety Set	125 (100)	121 (100)	246 (100)
Number of Patients Excluded from the Safety Set	0 (0.0)	0 (0.0)	0 (0.0)
Number of Patients in the PK Analysis Set	125 (100)	121 (100)	246 (100)
Number of Patients Excluded from the PK Analysis Set	0 (0.0)	0 (0.0)	0 (0.0)

Note: Percentages are based on the number of randomized patients in each column.

Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.

Results

A little bit of text appears here. A little bit of text appears here. **A little bit of text appears here.**
A little bit of text appears here.

Some text appears here.

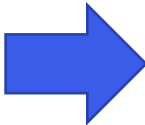
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here.

<Figure 1>

Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.

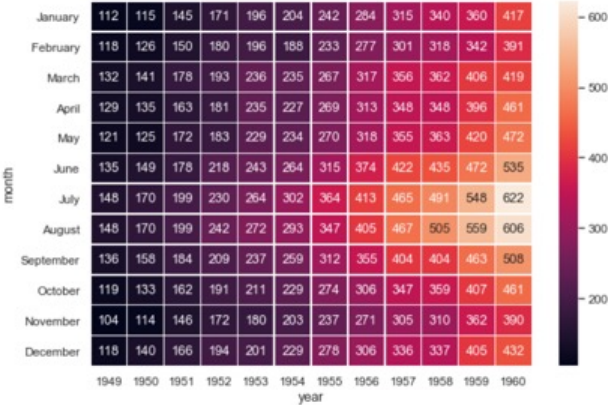
<Table 3>

A little bit of text appears here. A little bit of text appears here. **A little bit of text appears here.**
A little bit of text appears here.



Lots of text appear here. Lots of text appear here. Lots of text appear here. Lots of text appear here.
Lots of text appear here. Lots of text appear here.
Lots of text appear here.
Lots of text appear here.

Figure 1: EORTC QLQ-30 Fatigue Subscale: Percentage of Patients with a 10 Point Improvement from Baseline, by Visit (Full Analysis Set)



The plot will be similar to figure 14.2.2.2.1.
Note: EORTC QLQ-C30=European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire-Core 30 Scale.

Lots of text appear here.
Lots of text appear here.

Calculations

Let count how much time you need to update Clinical Study Report (CSR) for AstraZeneca trial 5077US/0049

Standard solution

1. update 70 Table and 10 Figures (5 minutes * 80 = 400 minutes = 7 hours)
2. Wait until programmers fix tables with cosmetic issues (4 hours)
3. Manual check (1 hour)

Automated solution

1. Run utility (1 minute)
2. Wait for compiling (80 objects * 30 seconds = 40 minutes)
3. Manual check (1 hour)

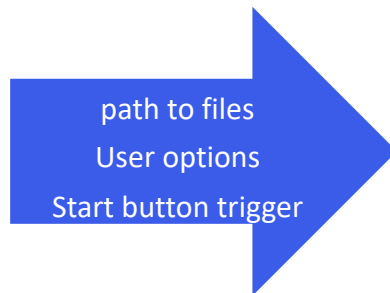
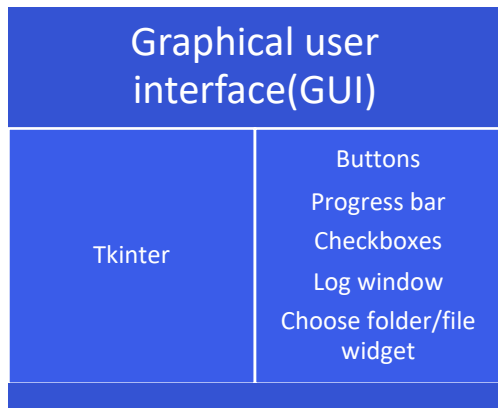
Benefits of using automatized approach

- Minimize action to update report (1 minute to start the utility and 30 seconds * (number of tables) to compile document)
- Reduce time for checking titles and footnotes.
- Using sas7bdat dataset as a source. You don't need to wait till programmers will fix cosmetic issues



Programming background

- Developed using Python 3.7 + Tkinter for a GUI
- 2 modules: doc_util (logic) and GUI(graphical user interface)



Intext logic libraries

Sas7bdat

Read data from sas7bdat files and Convert them to DataFrame object

docx

Python library for creating and updating Microsoft Word files

docx.xml

Create tables in Microsoft Word files.
Allows to specify table size, borders and fonts

Future development

- Run intext utility on the real project
- Add partial compiling of document to reduce compilation time. Update only new parts, instead all document
- Add more client specified features



<https://www.hospitallers.life/>



Swift

USD, Dollar

Receiver:

FEDIANOVYCH HANNA VITALIIIVNA

Recipient address:

51400, Ukraine, Dnipropetrovskyy region, city Pavlohrad, street Zakhidna, building 13

IBAN:

UA973052990262086400928825287

Recipient's bank:

JSC CB PRIVATBANK, 1D HRUSHEVSKOHO STR., KYIV, 01001, UKRAINE

SWIFT CODE/BIC:

PBANUA2X



Thanks for your attention

