



**Veristat TLF Packager (GUI):
An Interactive Tool for
Automated Creation of Combined PDFs
With Bookmarks &
Clickable Table of Contents**

Manivannan Mathialagan

Agenda

- ❖ Background & Motivation
- ❖ Challenges in Current TLF Packaging
- ❖ Solution Overview
- ❖ Key Features
- ❖ System Architecture & Implementation
- ❖ Results & Benefits
- ❖ Conclusion

Background & Motivation

- ❖ Clinical programming teams handle large volumes of Tables, Listings, and Figures (TLFs) across multiple deliverables (CSR, DSMB, DMC, DSUR, etc.).
- ❖ Packaging these outputs into a clean, ordered, bookmarked PDF is essential for review, regulatory submission, and quality control.
- ❖ Existing processes rely heavily on manual steps—reviewing titles, renaming files, ordering outputs, validating consistency, and merging PDFs.
- ❖ These manual workflows lead to inefficiencies, inconsistent output, increased QC burden, and risk of human error.

Background & Motivation

- ❖ A need emerged for an intuitive, automated, and user-friendly solution that could streamline TLF packaging, reduce effort, and produce a consistent, high-quality combined PDF every time.
- ❖ This motivated the development of an interactive GUI tool that simplifies the workflow, improves accuracy, and saves significant programming and QC time.

Challenges in Current TLF Packaging

- ❖ Manual title extraction from heterogeneous formats.
- ❖ Bookmarks missing or incorrectly named.
- ❖ Difficult to reorder or filter outputs before merging.
- ❖ No easy way to convert RTF/DOCX to PDF consistently.
- ❖ Packaging workflows differ across studies and programmers.
- ❖ High dependency on Word/PDF manual steps → inconsistent audit trail.

Solution Overview

- ❖ A standalone GUI EXE tool built on Python & PyQt5.
- ❖ Scans RTF, DOCX, and PDF files and auto-extracts Titles.
- ❖ Allows interactive review, filtering, and reordering.
- ❖ Converts RTF/DOCX → PDF via MS Word integration.
- ❖ Generates a single merged PDF with bookmarks + a clickable Table of Contents (TOC).

Key Features

- ❖ Automatic title extraction from RTF/DOCX/PDF.
- ❖ Search, filter, and sort TLFs interactively.
- ❖ Editable Bookmark field and Include/Exclude toggles.
- ❖ Move Up / Move Down reordering.
- ❖ TOC generation with wrapped, clickable links.
- ❖ Integrated PDF merger with hierarchical bookmarks.
- ❖ Log window with color-coded messages (Success, Warning, Error).

Key Features

- ❖ TOC checkbox to include/exclude TOC pages.
- ❖ Terminate button to stop ongoing processes.
- ❖ Outputs Excel export of selections and ordering.

Architecture & Workflow

- ❖ File Scan Layer:
 - ❖ Reads RTF (header/body), DOCX (paragraphs), PDF (page 1).
 - ❖ Extracts titles using custom parsing functions.
- ❖ GUI Interaction Layer:
 - ❖ PyQt5 table to display all outputs.
 - ❖ Checkbox, bookmark edit, search, and ordering controls.

Architecture & Workflow

❖ Conversion Layer:

- ❖ Uses MS Word automation (pywin32) to convert RTF/DOCX → PDF.
- ❖ Tracks converted paths for downstream merging.

❖ Packaging Layer:

- ❖ Merges PDFs using PyMuPDF.
- ❖ Generates TOC.
- ❖ Creates hierarchical bookmarks (Study → Outputs).

Architecture & Workflow

- ❖ Export Layer:

- ❖ Produces Excel detailing ordering, bookmarks, and metadata.

Technical Implementation

- ❖ Python 3, PyQt5, PyMuPDF (fitz), python-docx, pywin32.
- ❖ Multi-threading to prevent GUI freezing during heavy operations.
- ❖ Robust error handling and status logging.
- ❖ Executable packaged using PyInstaller for non-Python users.
- ❖ Handles filenames with special characters and long paths.
- ❖ Supports custom TOC formatting and bookmark nesting.

Example Usage - Screenshots

Veristat TLF Packager

Veristat TLF Packager

• For any queries, issues faced, or if you have ideas for improvement, please contact Manivannan (manivannan.mathialagan@veristat.com)

Search: Include TOC page Selected: 0 RTF | 2 DOCX | 31 PDF Bookmark text is editable (double-click), and Include column can be checked/unchecked.

	Include	Type	Filename	Title 1	Title 2	Title 3	Bookmark
1	<input checked="" type="checkbox"/>	DOCX	t14_1_1_1_1.docx	Table 14.1.1.1.1	Patient Disposition	Intent to Treat Population	Table 14.1.1.1.1: Patient ...
2	<input checked="" type="checkbox"/>	PDF	t14_1_1_1_2.pdf	Table 14.1.1.1.2	Patient Disposition	Safety Population	Table 14.1.1.1.2: Patient ...
3	<input checked="" type="checkbox"/>	PDF	t14_1_1_1_3.pdf	Table 14.1.1.1.3	Patient Disposition – By ...	Intent to Treat Population	Table 14.1.1.1.3: Patient ...
4	<input checked="" type="checkbox"/>	PDF	t14_1_1_1_4.pdf	Table 14.1.1.1.4	Patient Disposition – By ...	Safety Population	Table 14.1.1.1.4: Patient ...
5	<input checked="" type="checkbox"/>	PDF	t14_1_3_1_1.pdf	Table 14.1.3.1.1	Demographics and Diagnosi...	Intent to Treat Population	Table 14.1.3.1.1: Demographi...
6	<input checked="" type="checkbox"/>	PDF	t14_1_3_1_2.pdf	Table 14.1.3.1.2	Demographics and Diagnosi...	Safety Population	Table 14.1.3.1.2: Demographi...
7	<input checked="" type="checkbox"/>	PDF	t14_1_3_1_3.pdf	Table 14.1.3.1.3	Demographics and Diagnosi...	Intent to Treat Population	Table 14.1.3.1.3: Demographi...
8	<input checked="" type="checkbox"/>	PDF	t14_1_3_1_4.pdf	Table 14.1.3.1.4	Demographics and Diagnosi...	Safety Population	Table 14.1.3.1.4: Demographi...
9	<input checked="" type="checkbox"/>	PDF	t14_1_3_2_1.pdf	Table 14.1.3.2.1	Baseline Clinical Presentation	Intent to Treat Population	Table 14.1.3.2.1: Baseline ...
10	<input checked="" type="checkbox"/>	PDF	t14_1_3_2_2.pdf	Table 14.1.3.2.2	Baseline Clinical Presentation	Safety Population	Table 14.1.3.2.2: Baseline ...
11	<input checked="" type="checkbox"/>	PDF	t14_1_3_2_3.pdf	Table 14.1.3.2.3	Baseline Clinical Presentatio...	Intent to Treat Population	Table 14.1.3.2.3: Baseline ...
12	<input checked="" type="checkbox"/>	PDF	t14_1_3_2_4.pdf	Table 14.1.3.2.4	Baseline Clinical Presentatio...	Safety Population	Table 14.1.3.2.4: Baseline ...

Output PDF Name:

Log:
Auto-refresh is enabled (every 3 seconds). Any new file will be auto-detected.

Example Usage - Screenshots

Veristat TLF Packager

- □ ×

Veristat TLF Packager

• For any queries, issues faced, or if you have ideas for improvement, please contact Manivannan (manivannan.mathialagan@veristat.com)

 Search:
 Include TOC page

Selected: 0 RTF | 2 DOCX | 32 PDF

Bookmark text is editable (double-click), and Include column can be checked/unchecked.

	Include	Type	Filename	Title 1	Title 2	Title 3	Bookmark
1	<input checked="" type="checkbox"/>	DOCX	t14_1_1_1_1.docx	Table 14.1.1.1.1	Patient Disposition	Intent to Treat Population	Table 14.1.1.1.1: Patient ...
2	<input checked="" type="checkbox"/>	PDF	t14_1_1_1_2.pdf	Table 14.1.1.1.2	Patient Disposition	Safety Population	Table 14.1.1.1.2: Patient ...
3	<input checked="" type="checkbox"/>	PDF	t14_1_1_1_3.pdf	Table 14.1.1.1.3	Patient Disposition – By ...	Intent to Treat Population	Table 14.1.1.1.3: Patient ...
4	<input checked="" type="checkbox"/>	PDF	t14_1_1_1_4.pdf	Table 14.1.1.1.4	Patient Disposition – By ...	Safety Population	Table 14.1.1.1.4: Patient ...
5	<input checked="" type="checkbox"/>	PDF				Intent to Treat Population	Table 14.1.3.1.1: Demographi...
6	<input checked="" type="checkbox"/>	PDF				Safety Population	Table 14.1.3.1.2: Demographi...
7	<input checked="" type="checkbox"/>	PDF				Intent to Treat Population	Table 14.1.3.1.3: Demographi...
8	<input checked="" type="checkbox"/>	PDF				Safety Population	Table 14.1.3.1.4: Demographi...
9	<input checked="" type="checkbox"/>	PDF				Intent to Treat Population	Table 14.1.3.2.1: Baseline ...
10	<input checked="" type="checkbox"/>	PDF				Safety Population	Table 14.1.3.2.2: Baseline ...
11	<input checked="" type="checkbox"/>	PDF	t14_1_3_2_3.pdf	Table 14.1.3.2.3	Baseline Clinical Presentatio...	Intent to Treat Population	Table 14.1.3.2.3: Baseline ...
12	<input checked="" type="checkbox"/>	PDF	t14_1_3_2_4.pdf	Table 14.1.3.2.4	Baseline Clinical Presentatio...	Safety Population	Table 14.1.3.2.4: Baseline ...

Done

Output PDF generated:
 P:/Projects/Veristat/BSP Demo Project/BSP/Output/Development TLFs/
 tables_shells\TLFs_Merged_20251206_T1619.pdf

OK

Browse Folder

Select All

Export to Excel

Move Up

Move Down

Convert to PDF

Pack & TOC

Terminate

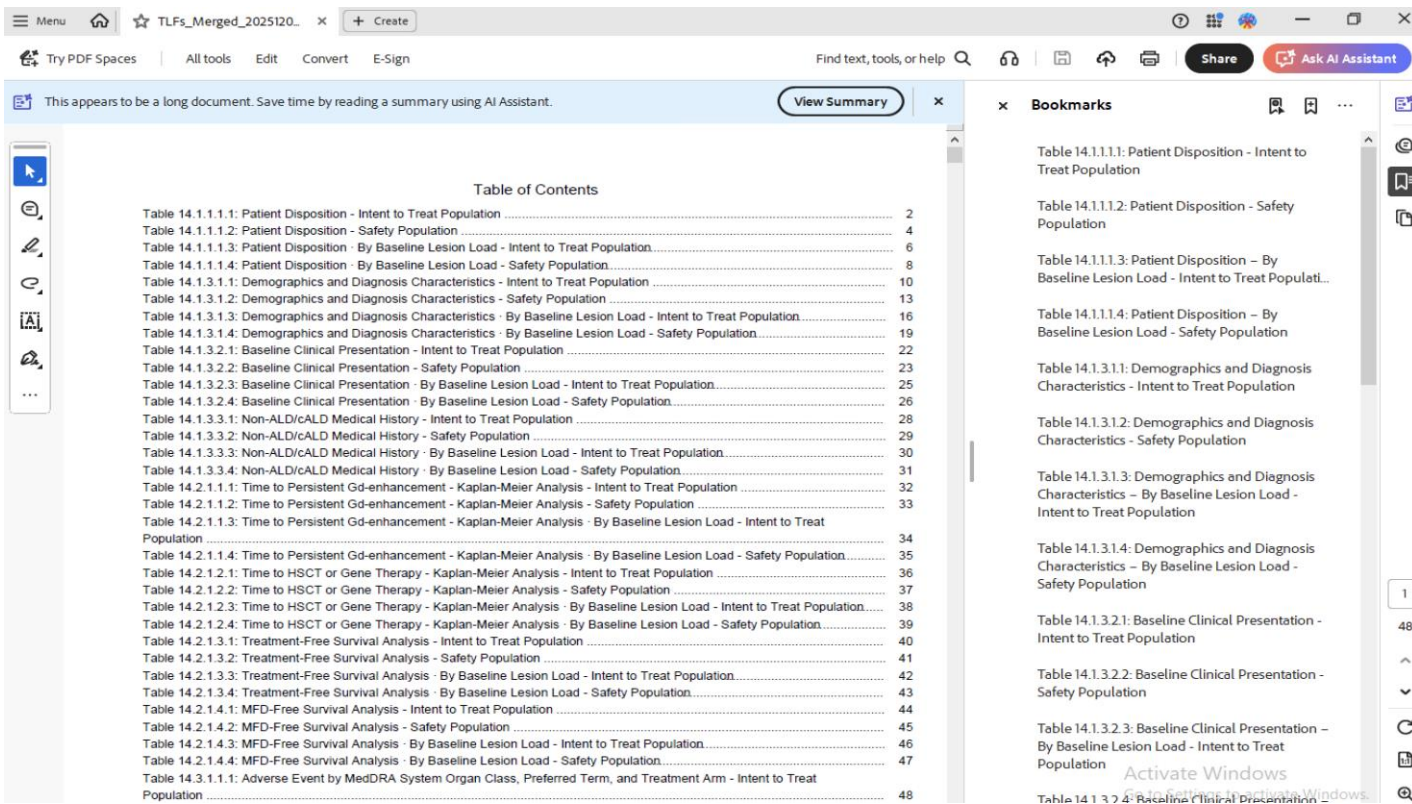
 Output PDF Name:

Log

```

[29/33] Adding: t14_2_1_4_1.pdf (PDF)
[30/33] Adding: t14_2_1_4_2.pdf (PDF)
[31/33] Adding: t14_2_1_4_3.pdf (PDF)
[32/33] Adding: t14_2_1_4_4.pdf (PDF)
[33/33] Adding: t14_3.1.1.1-rt-ae-itt.docx (DOCX)
[Merging] Inserting TOC and bookmarks...
[DONE] PDF created: P:/Projects/Veristat/BSP Demo Project/BSP/Output/Development TLFs/tables_shells\TLFs_Merged_20251206_T1619.pdf
  
```

Example Usage - Screenshots



The screenshot shows a web-based PDF viewer interface. At the top, there is a navigation bar with a 'Menu' icon, a home icon, and a browser tab titled 'TLFs_Merged_2025120...'. Below this is a toolbar with options like 'Try PDF Spaces', 'All tools', 'Edit', 'Convert', and 'E-Sign'. A search bar contains the text 'Find text, tools, or help'. To the right of the search bar are icons for 'Share' and 'Ask AI Assistant'. Below the toolbar, a blue notification bar states: 'This appears to be a long document. Save time by reading a summary using AI Assistant.' with a 'View Summary' button. The main content area displays a 'Table of Contents' with a list of sections and their corresponding page numbers. A sidebar on the right is titled 'Bookmarks' and lists several table entries with their page numbers. At the bottom of the sidebar, there are navigation controls including a page number '1', a list of page numbers '48', and a search icon.

Table of Contents

Table 14.1.1.1.1: Patient Disposition - Intent to Treat Population	2
Table 14.1.1.1.2: Patient Disposition - Safety Population	4
Table 14.1.1.1.3: Patient Disposition - By Baseline Lesion Load - Intent to Treat Population	6
Table 14.1.1.1.4: Patient Disposition - By Baseline Lesion Load - Safety Population	8
Table 14.1.3.1.1: Demographics and Diagnosis Characteristics - Intent to Treat Population	10
Table 14.1.3.1.2: Demographics and Diagnosis Characteristics - Safety Population	13
Table 14.1.3.1.3: Demographics and Diagnosis Characteristics - By Baseline Lesion Load - Intent to Treat Population	16
Table 14.1.3.1.4: Demographics and Diagnosis Characteristics - By Baseline Lesion Load - Safety Population	19
Table 14.1.3.2.1: Baseline Clinical Presentation - Intent to Treat Population	22
Table 14.1.3.2.2: Baseline Clinical Presentation - Safety Population	23
Table 14.1.3.2.3: Baseline Clinical Presentation - By Baseline Lesion Load - Intent to Treat Population	25
Table 14.1.3.2.4: Baseline Clinical Presentation - By Baseline Lesion Load - Safety Population	26
Table 14.1.3.3.1: Non-ALD/rCALD Medical History - Intent to Treat Population	28
Table 14.1.3.3.2: Non-ALD/rCALD Medical History - Safety Population	29
Table 14.1.3.3.3: Non-ALD/rCALD Medical History - By Baseline Lesion Load - Intent to Treat Population	30
Table 14.1.3.3.4: Non-ALD/rCALD Medical History - By Baseline Lesion Load - Safety Population	31
Table 14.2.1.1.1: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - Intent to Treat Population	32
Table 14.2.1.1.2: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - Safety Population	33
Table 14.2.1.1.3: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - By Baseline Lesion Load - Intent to Treat Population	34
Table 14.2.1.1.4: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - By Baseline Lesion Load - Safety Population	35
Table 14.2.1.2.1: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - Intent to Treat Population	36
Table 14.2.1.2.2: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - Safety Population	37
Table 14.2.1.2.3: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - By Baseline Lesion Load - Intent to Treat Population	38
Table 14.2.1.2.4: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - By Baseline Lesion Load - Safety Population	39
Table 14.2.1.3.1: Treatment-Free Survival Analysis - Intent to Treat Population	40
Table 14.2.1.3.2: Treatment-Free Survival Analysis - Safety Population	41
Table 14.2.1.3.3: Treatment-Free Survival Analysis - By Baseline Lesion Load - Intent to Treat Population	42
Table 14.2.1.3.4: Treatment-Free Survival Analysis - By Baseline Lesion Load - Safety Population	43
Table 14.2.1.4.1: MFD-Free Survival Analysis - Intent to Treat Population	44
Table 14.2.1.4.2: MFD-Free Survival Analysis - Safety Population	45
Table 14.2.1.4.3: MFD-Free Survival Analysis - By Baseline Lesion Load - Intent to Treat Population	46
Table 14.2.1.4.4: MFD-Free Survival Analysis - By Baseline Lesion Load - Safety Population	47
Table 14.3.1.1.1: Adverse Event by MedDRA System Organ Class, Preferred Term, and Treatment Arm - Intent to Treat Population	48

Bookmarks

- Table 14.1.1.1.1: Patient Disposition - Intent to Treat Population
- Table 14.1.1.1.2: Patient Disposition - Safety Population
- Table 14.1.1.1.3: Patient Disposition - By Baseline Lesion Load - Intent to Treat Populat...
- Table 14.1.1.1.4: Patient Disposition - By Baseline Lesion Load - Safety Population
- Table 14.1.3.1.1: Demographics and Diagnosis Characteristics - Intent to Treat Population
- Table 14.1.3.1.2: Demographics and Diagnosis Characteristics - Safety Population
- Table 14.1.3.1.3: Demographics and Diagnosis Characteristics - By Baseline Lesion Load - Intent to Treat Population
- Table 14.1.3.1.4: Demographics and Diagnosis Characteristics - By Baseline Lesion Load - Safety Population
- Table 14.1.3.2.1: Baseline Clinical Presentation - Intent to Treat Population
- Table 14.1.3.2.2: Baseline Clinical Presentation - Safety Population
- Table 14.1.3.2.3: Baseline Clinical Presentation - By Baseline Lesion Load - Intent to Treat Population
- Table 14.1.3.2.4: Baseline Clinical Presentation - By Baseline Lesion Load - Safety Population
- Table 14.1.3.3.1: Non-ALD/rCALD Medical History - Intent to Treat Population
- Table 14.1.3.3.2: Non-ALD/rCALD Medical History - Safety Population
- Table 14.1.3.3.3: Non-ALD/rCALD Medical History - By Baseline Lesion Load - Intent to Treat Population
- Table 14.1.3.3.4: Non-ALD/rCALD Medical History - By Baseline Lesion Load - Safety Population
- Table 14.2.1.1.1: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - Intent to Treat Population
- Table 14.2.1.1.2: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - Safety Population
- Table 14.2.1.1.3: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - By Baseline Lesion Load - Intent to Treat Population
- Table 14.2.1.1.4: Time to Persistent Gd-enhancement - Kaplan-Meier Analysis - By Baseline Lesion Load - Safety Population
- Table 14.2.1.2.1: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - Intent to Treat Population
- Table 14.2.1.2.2: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - Safety Population
- Table 14.2.1.2.3: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - By Baseline Lesion Load - Intent to Treat Population
- Table 14.2.1.2.4: Time to HSCT or Gene Therapy - Kaplan-Meier Analysis - By Baseline Lesion Load - Safety Population
- Table 14.2.1.3.1: Treatment-Free Survival Analysis - Intent to Treat Population
- Table 14.2.1.3.2: Treatment-Free Survival Analysis - Safety Population
- Table 14.2.1.3.3: Treatment-Free Survival Analysis - By Baseline Lesion Load - Intent to Treat Population
- Table 14.2.1.3.4: Treatment-Free Survival Analysis - By Baseline Lesion Load - Safety Population
- Table 14.2.1.4.1: MFD-Free Survival Analysis - Intent to Treat Population
- Table 14.2.1.4.2: MFD-Free Survival Analysis - Safety Population
- Table 14.2.1.4.3: MFD-Free Survival Analysis - By Baseline Lesion Load - Intent to Treat Population
- Table 14.2.1.4.4: MFD-Free Survival Analysis - By Baseline Lesion Load - Safety Population
- Table 14.3.1.1.1: Adverse Event by MedDRA System Organ Class, Preferred Term, and Treatment Arm - Intent to Treat Population

Results & Benefits

- ❖ ~80% reduction in packaging time per study.
- ❖ Eliminates manual bookmark creation.
- ❖ Ensures consistent naming and ordering across deliverables.
- ❖ Improves QC transparency with exportable Excel logs.
- ❖ Standardized workflow across programmers and studies.
- ❖ Reduces reliance on Word/PDF manual steps.
- ❖ Highly scalable for Phase 1–3 studies and DSMB/DSUR packages.

Use Cases

- ❖ CSR TLF Packaging
- ❖ DSMB Deliverable consolidation
- ❖ Interim Analysis packages
- ❖ Validation package creation for QC
- ❖ Ad-hoc statistical reports requiring bookmarked PDFs

Conclusion

- ❖ The GUI-based TLF Packager modernizes the TLF consolidation workflow.
- ❖ Provides a fast, consistent, and user-friendly solution across studies.
- ❖ Reduces manual effort, minimizes errors, and improves traceability.
- ❖ Designed for both programmers and QC reviewers.

Queries

