phuse·share
connect·share·advance

Single Day Event

Overcoming Challenges in Automating Report Creation for Clinical Trials Using NLP & DL

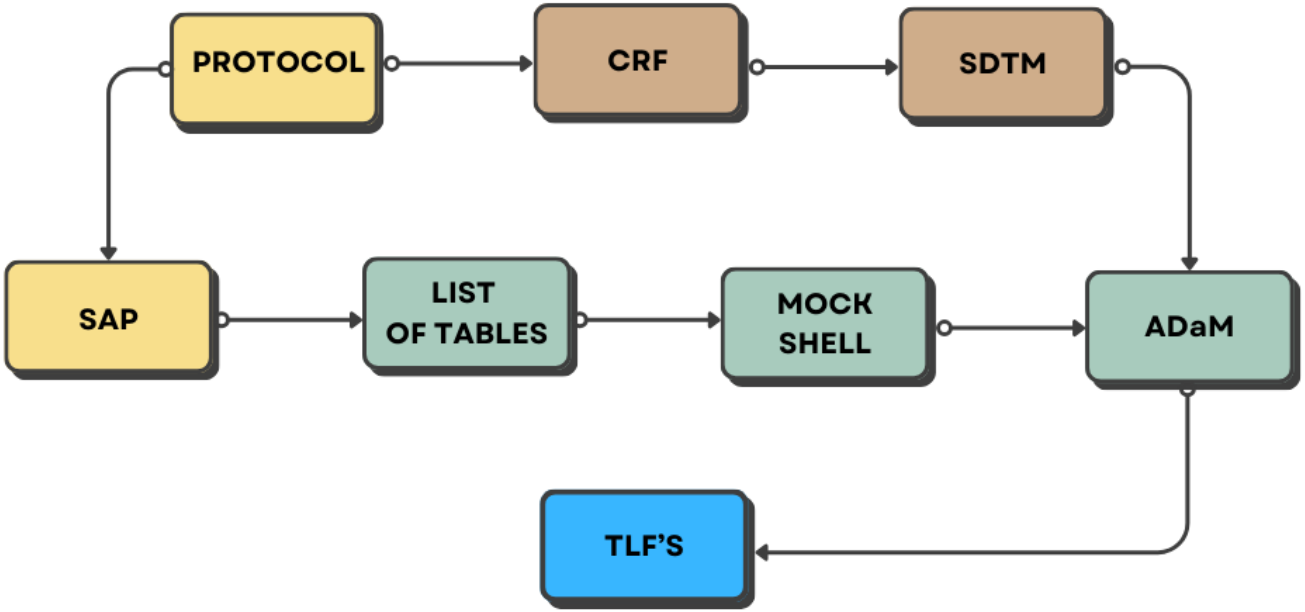
Surabhi Sundar,
Karthik Palani
Pfizer

phuse.global

Disclaimer

We confirm that, the opinions and thoughts discussed in this presentation are subject to our own independent views and are not subject to the opinions of the organization that we represent.

Report Generation Process: A Glimpse



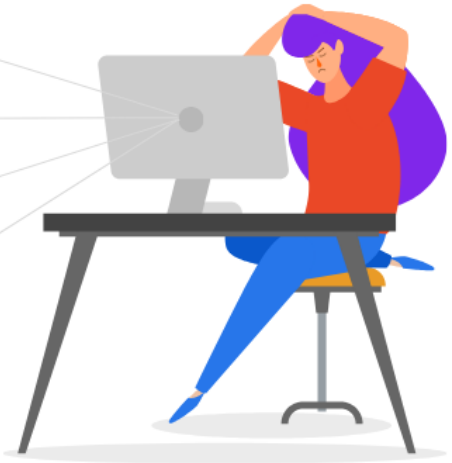
Report Generation Process: A Glimpse

Enhance Effectiveness 01

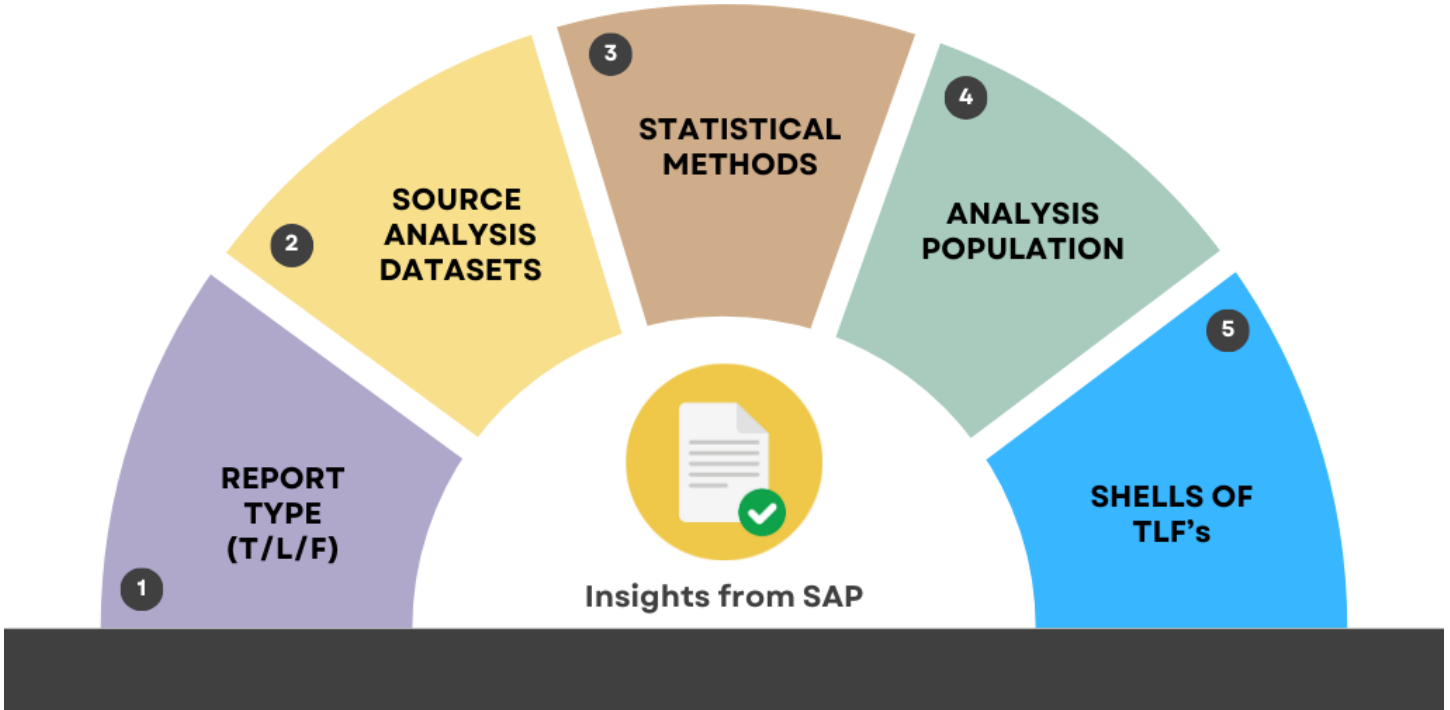
Streamline the Process 02

Reduce Dependency 03

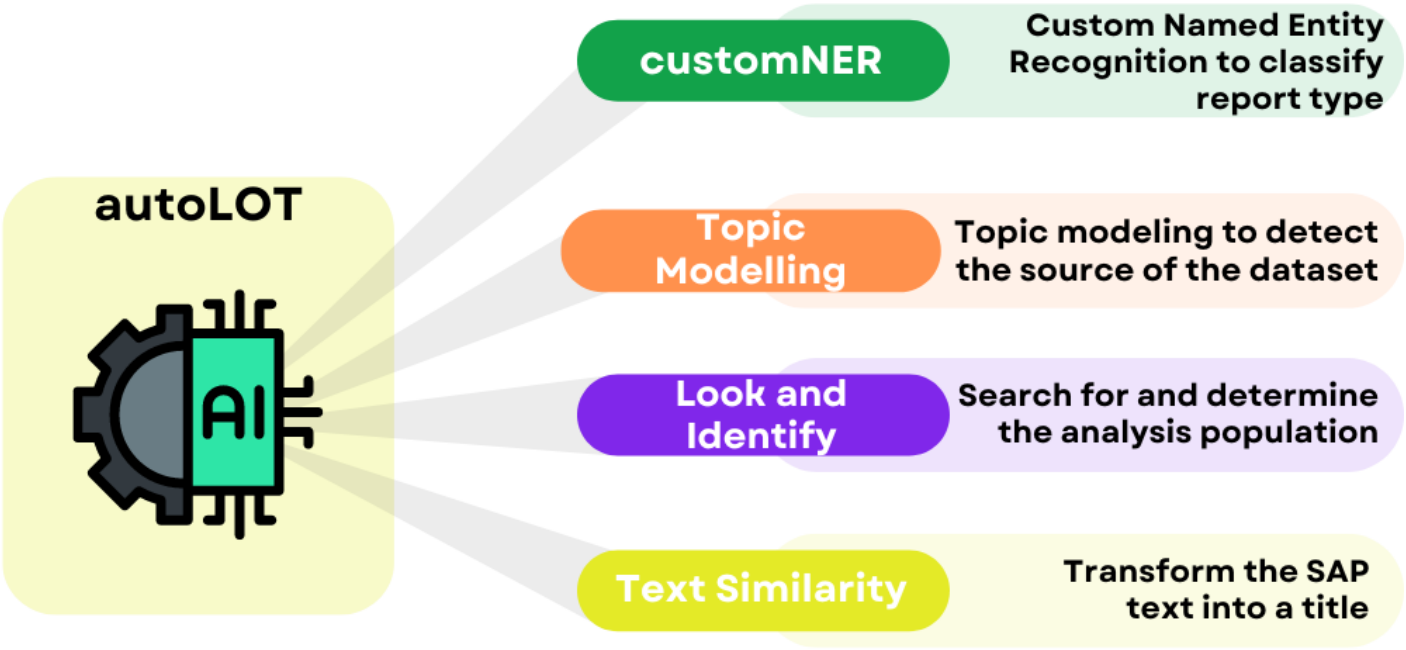
Time and Resources 04



Revamping Report Creation Procedure



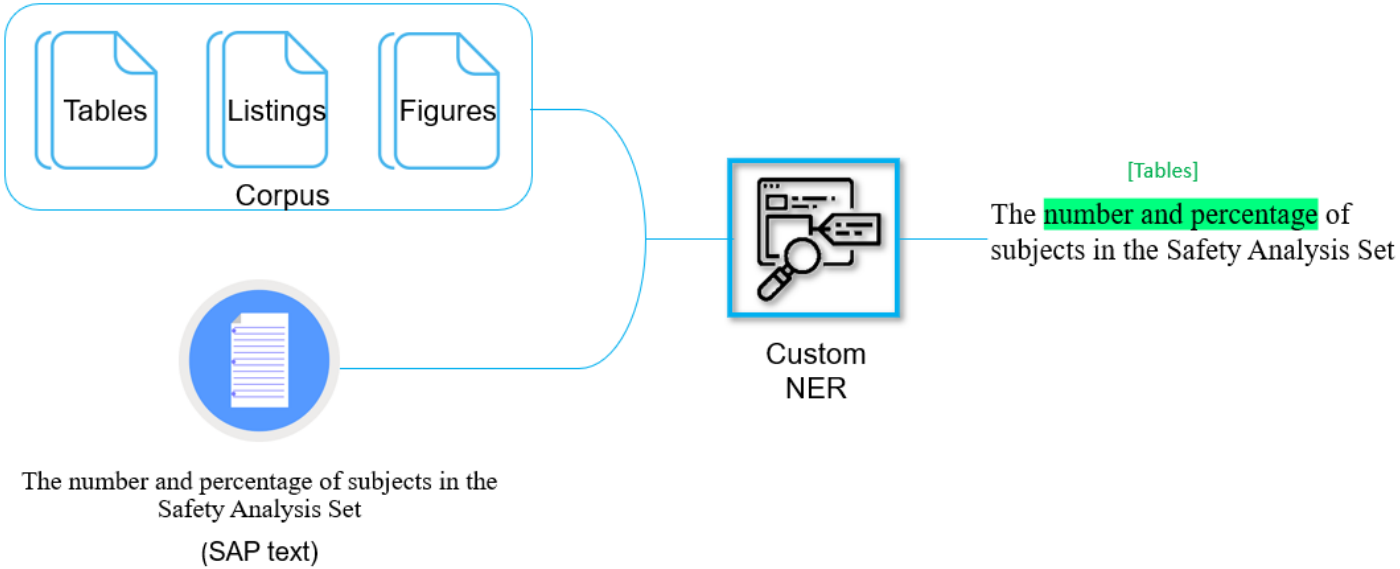
autoLOT Components



autoLOT Components

customNER — **Custom Named Entity Recognition to classify report type**

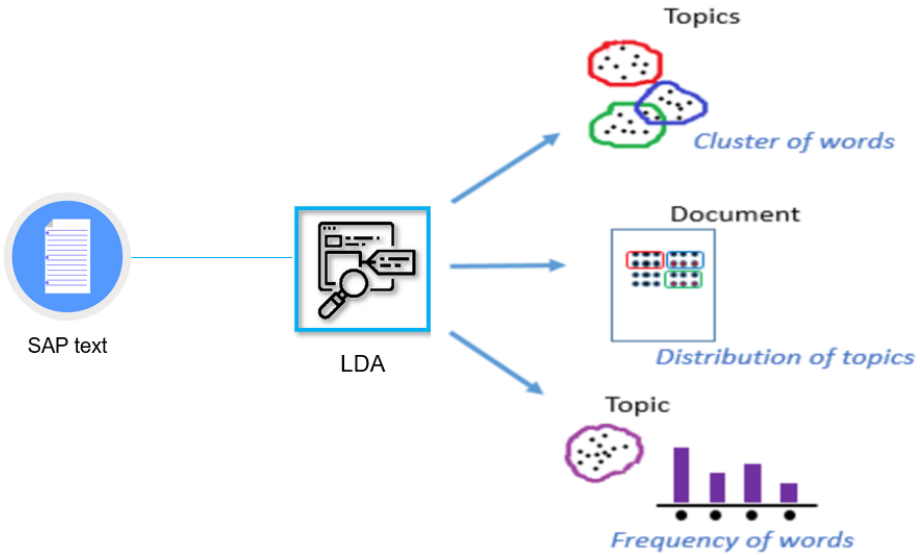
NER is an information extraction technique to identify and classify named entities in text



autoLOT Components

Topic Modelling — **Topic modeling to detect the source of the dataset**

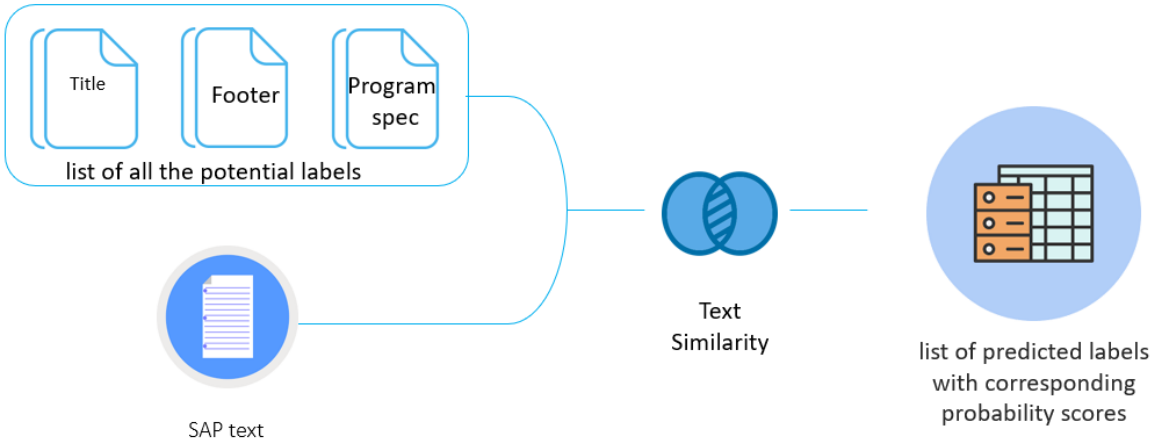
Topic Modelling is a text mining technique that analyzes the semantic patterns in text to make sense of unstructured data without the need for predefined training data.



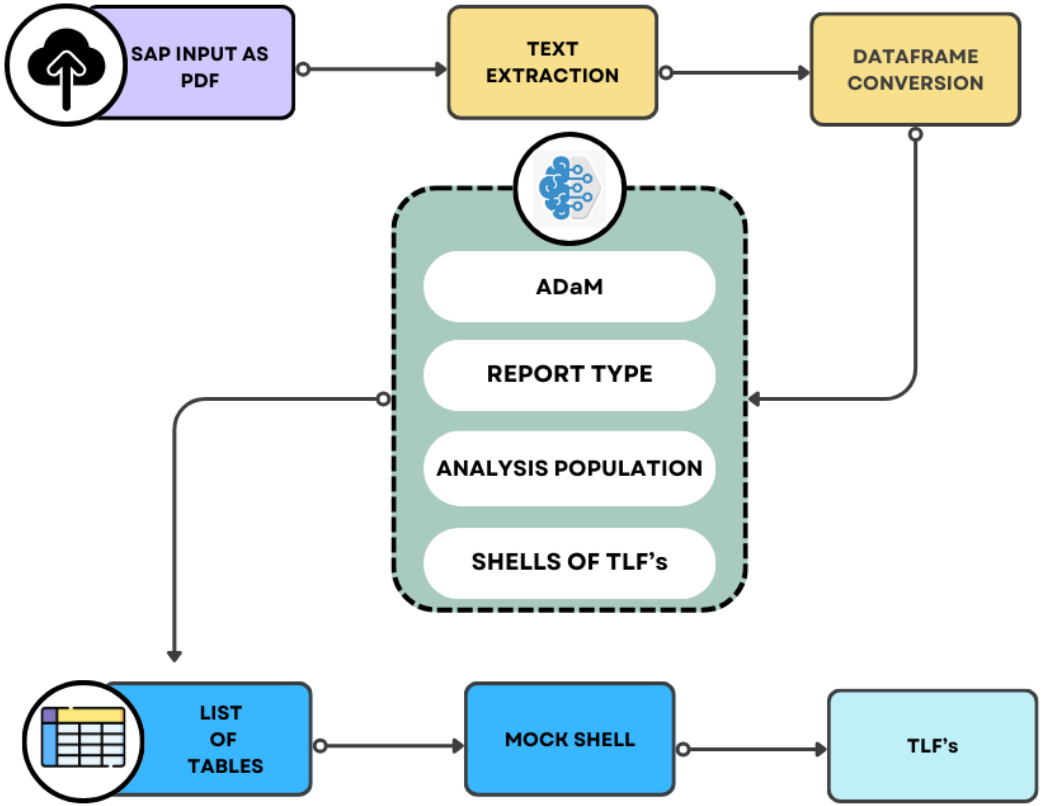
autoLOT Components

Text Similarity — **Transform the SAP text into a title**

Text Similarity is the process of comparing a piece of text with another and finding the similarity between them.

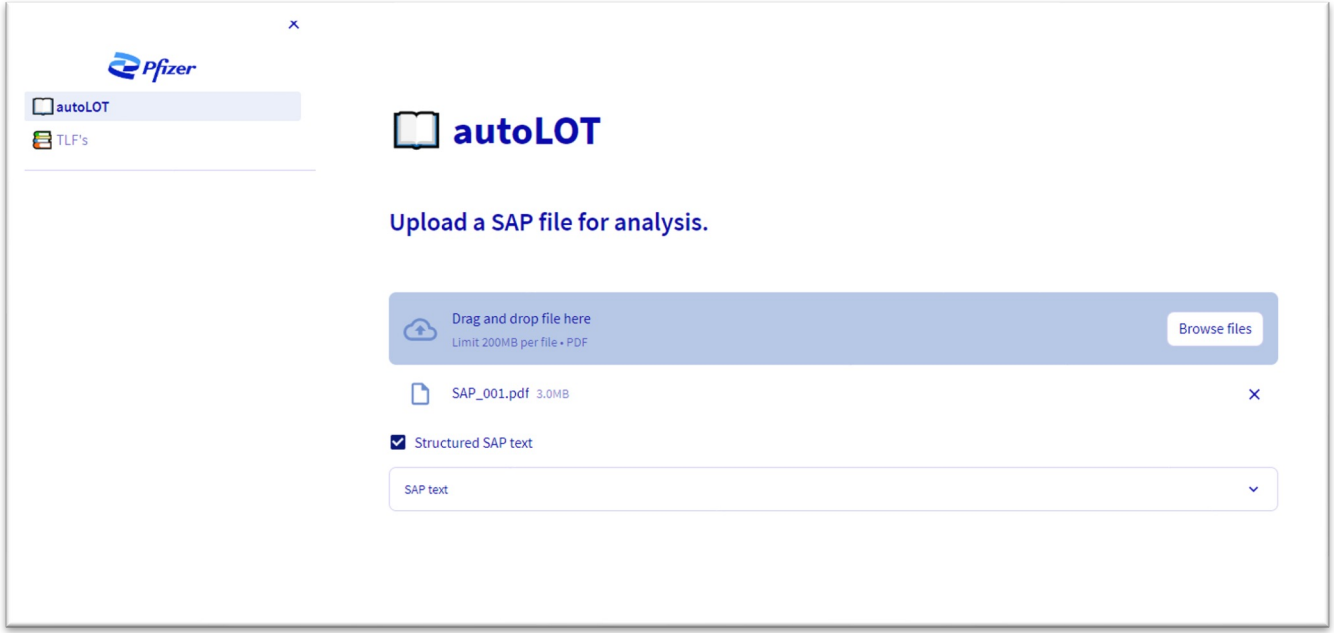


Discover the Functionality



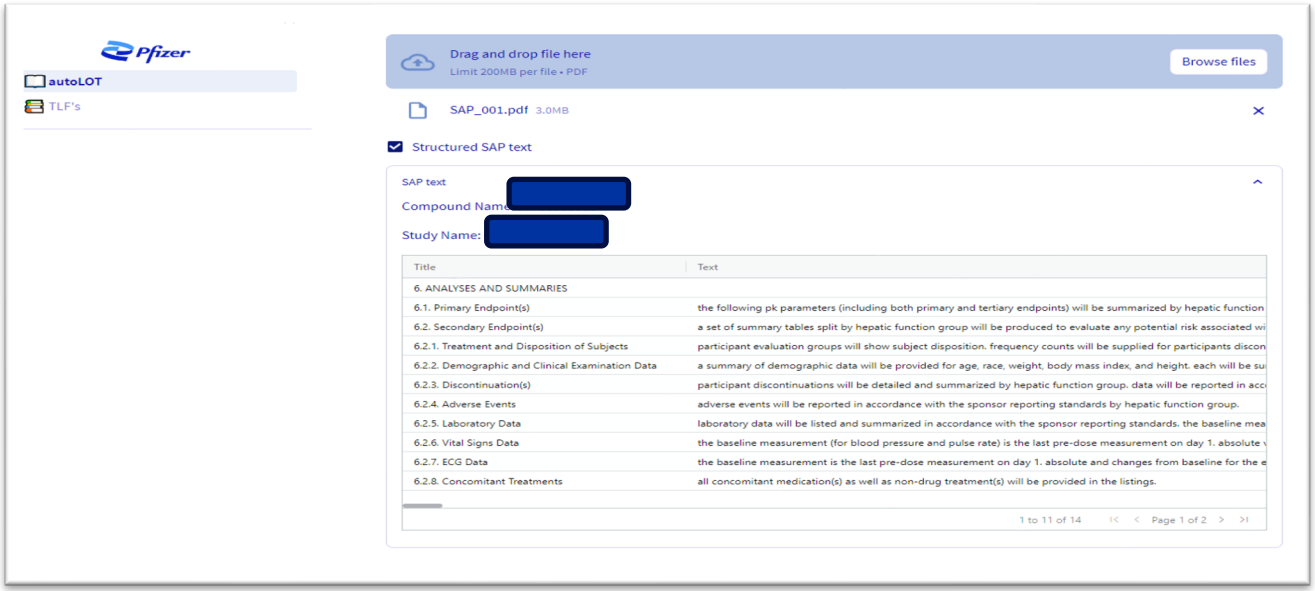
autoLOT: An Illustrative Showcase

Extracting text from SAP



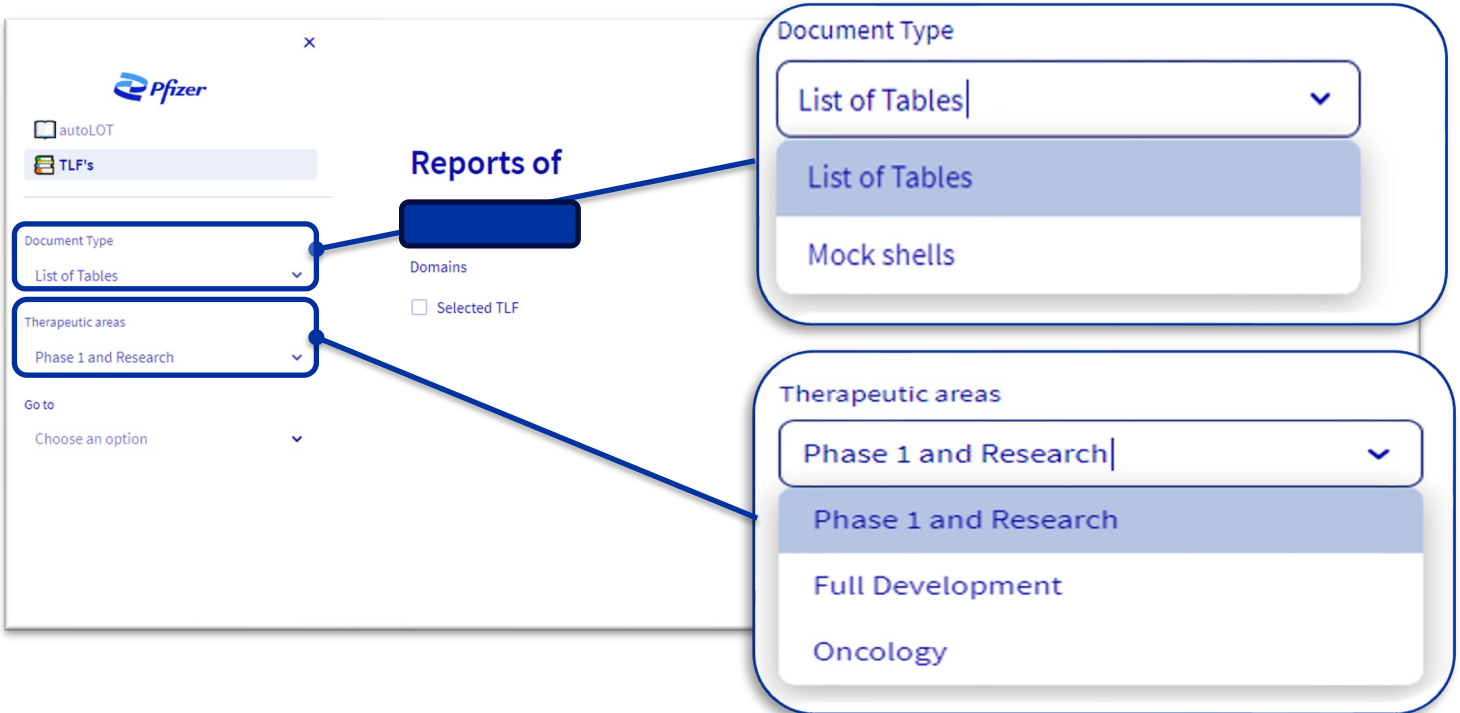
autoLOT: An Illustrative Showcase

Transform unstructured text data into a structured data frame



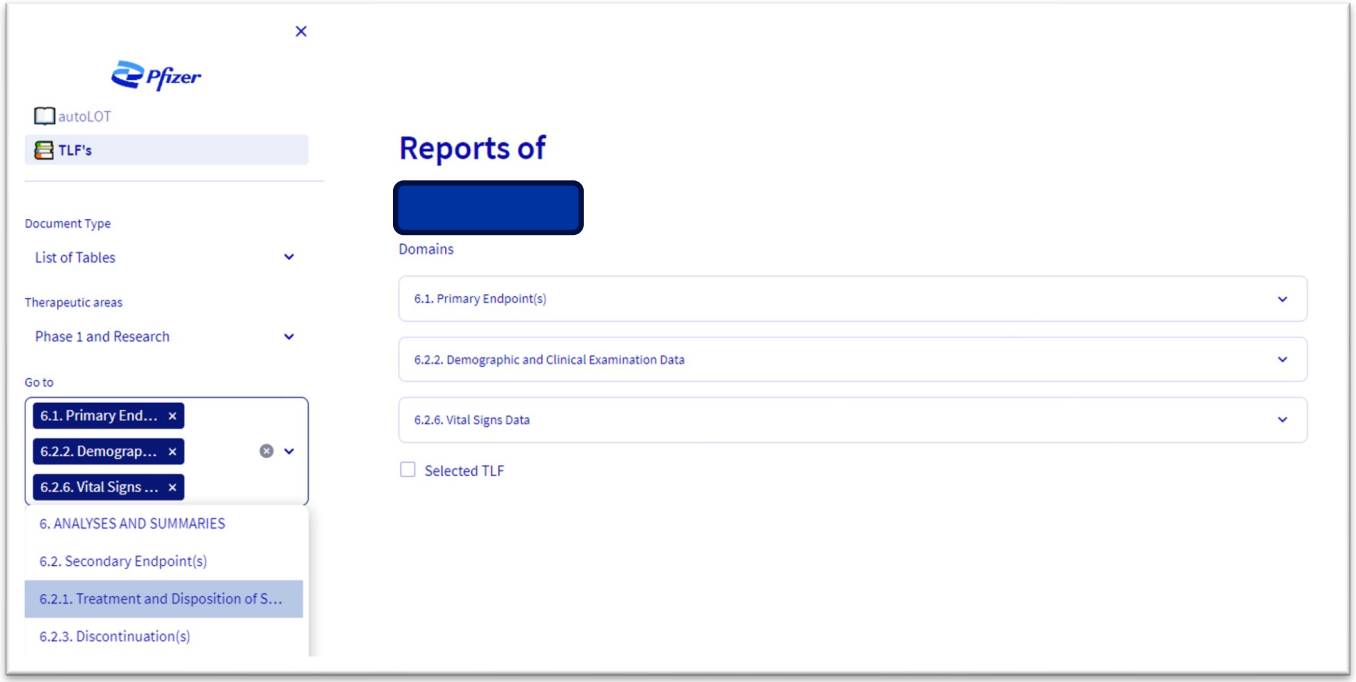
autoLOT: An Illustrative Showcase

Widgets to choose from



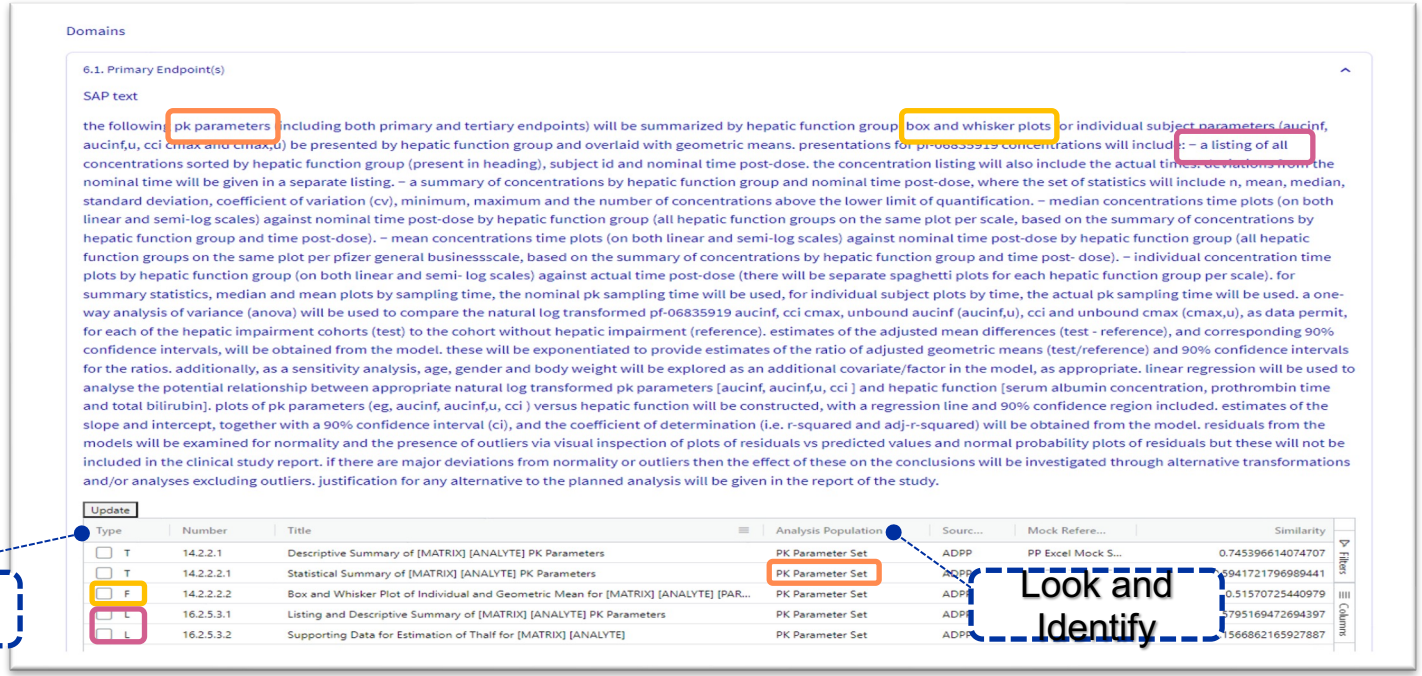
autoLOT: An Illustrative Showcase

Breaking apart into specific domains



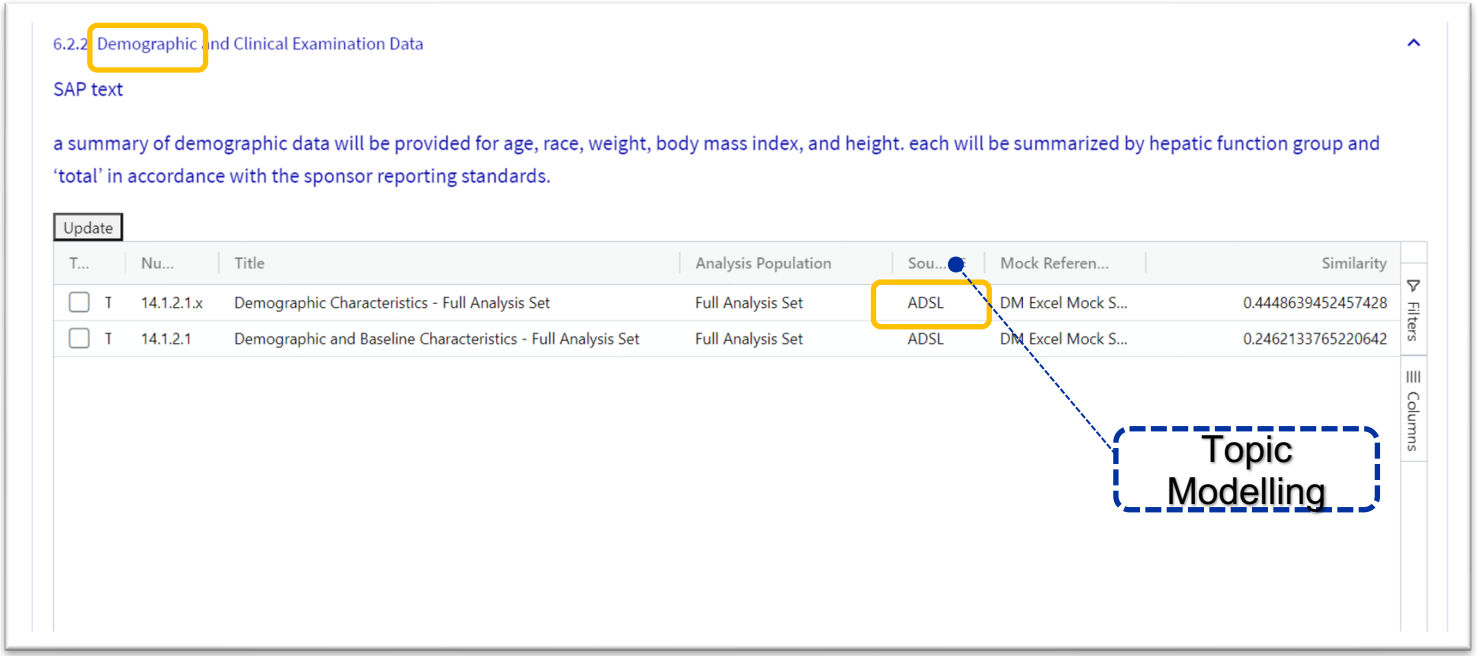
autoLOT: An Illustrative Showcase

Analysis of specific domains

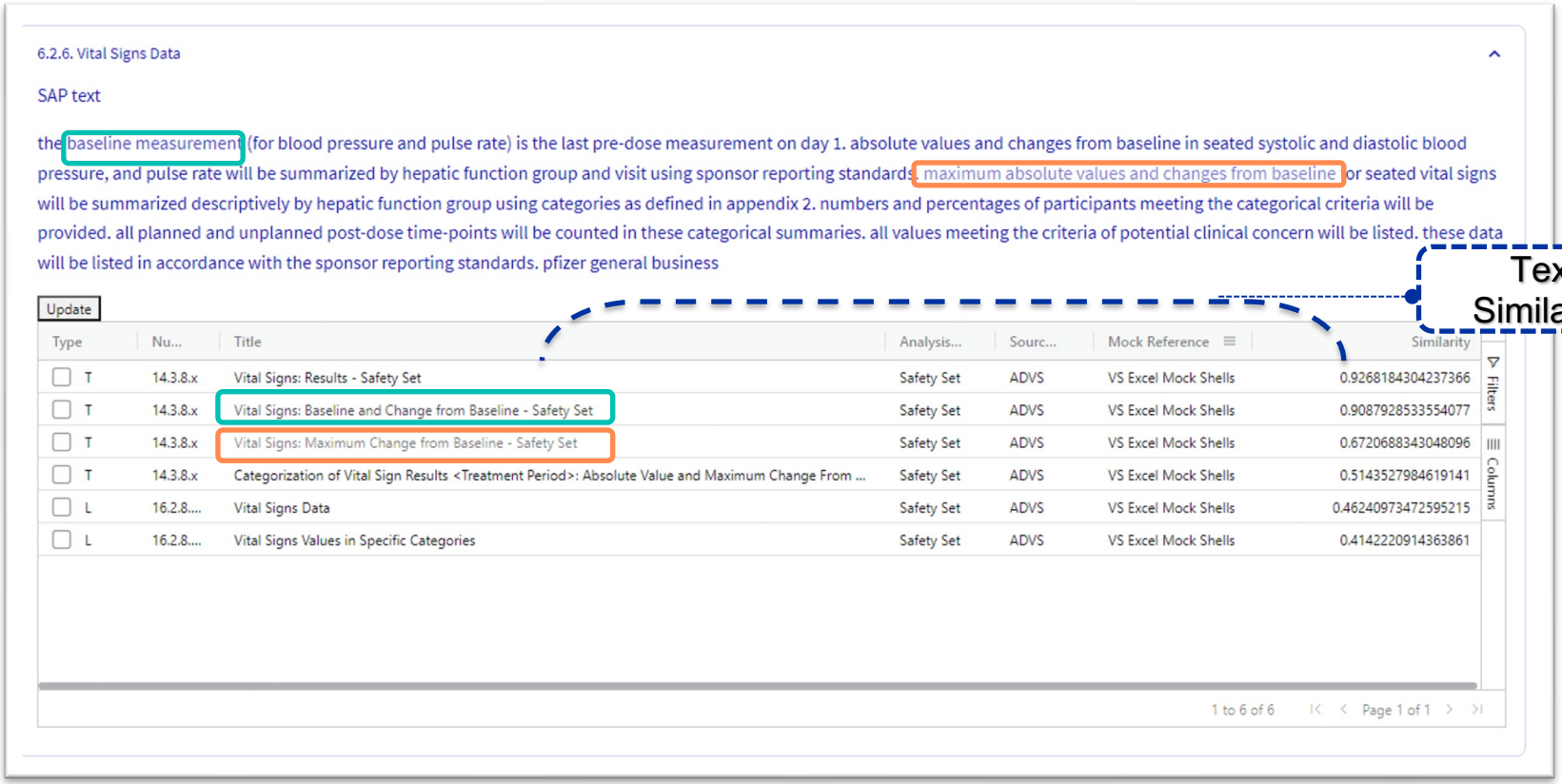


autoLOT: An Illustrative Showcase

Analysis of specific domains

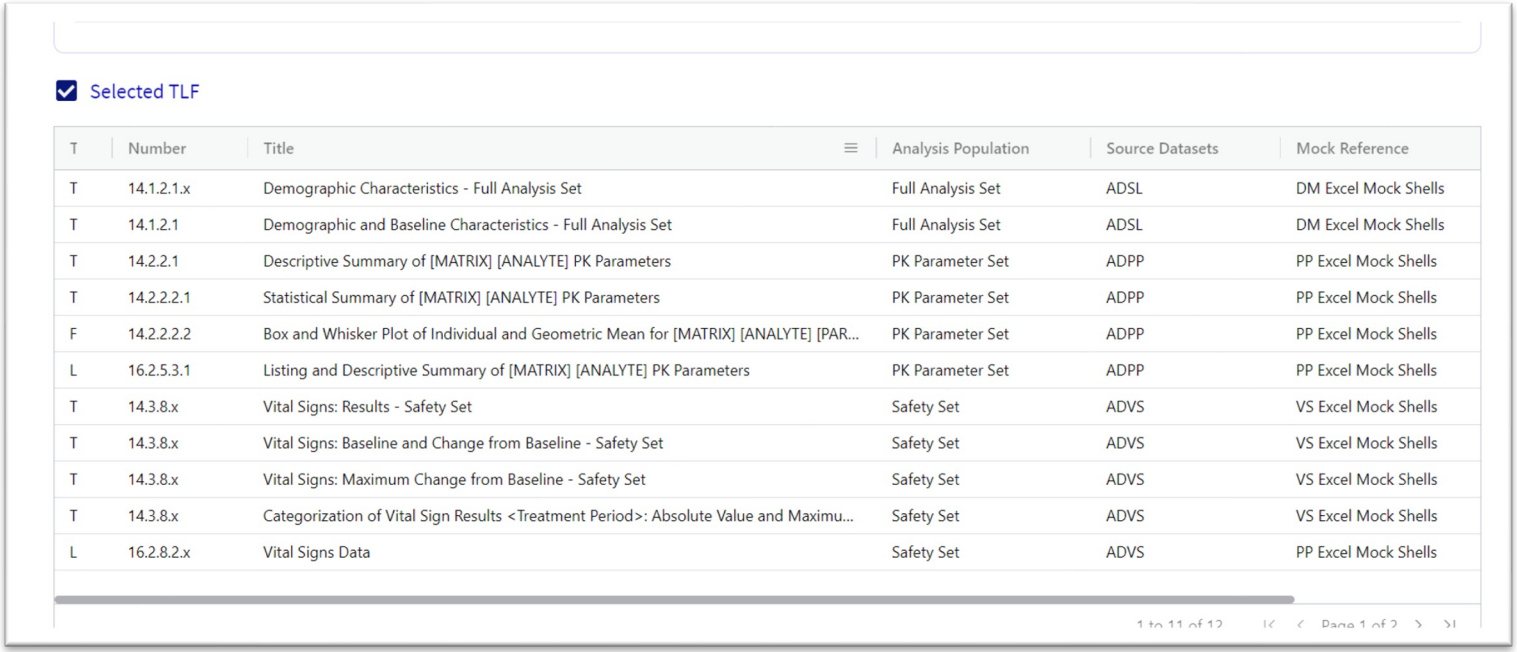


autoLOT: An Illustrative Showcase



autoLOT: An Illustrative Showcase

Consolidating all the required TLFs



Selected TLF

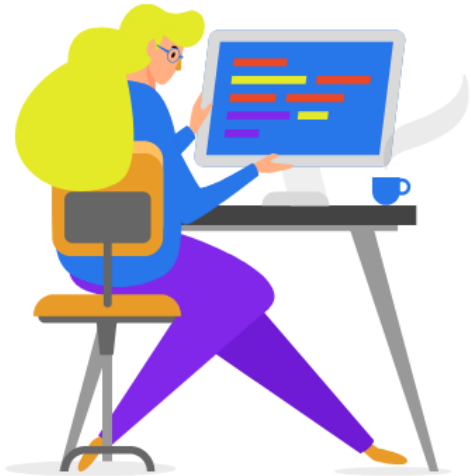
T	Number	Title	Analysis Population	Source Datasets	Mock Reference
T	14.1.2.1.x	Demographic Characteristics - Full Analysis Set	Full Analysis Set	ADSL	DM Excel Mock Shells
T	14.1.2.1	Demographic and Baseline Characteristics - Full Analysis Set	Full Analysis Set	ADSL	DM Excel Mock Shells
T	14.2.2.1	Descriptive Summary of [MATRIX] [ANALYTE] PK Parameters	PK Parameter Set	ADPP	PP Excel Mock Shells
T	14.2.2.2.1	Statistical Summary of [MATRIX] [ANALYTE] PK Parameters	PK Parameter Set	ADPP	PP Excel Mock Shells
F	14.2.2.2.2	Box and Whisker Plot of Individual and Geometric Mean for [MATRIX] [ANALYTE] [PAR...	PK Parameter Set	ADPP	PP Excel Mock Shells
L	16.2.5.3.1	Listing and Descriptive Summary of [MATRIX] [ANALYTE] PK Parameters	PK Parameter Set	ADPP	PP Excel Mock Shells
T	14.3.8.x	Vital Signs: Results - Safety Set	Safety Set	ADVS	VS Excel Mock Shells
T	14.3.8.x	Vital Signs: Baseline and Change from Baseline - Safety Set	Safety Set	ADVS	VS Excel Mock Shells
T	14.3.8.x	Vital Signs: Maximum Change from Baseline - Safety Set	Safety Set	ADVS	VS Excel Mock Shells
T	14.3.8.x	Categorization of Vital Sign Results <Treatment Period>: Absolute Value and Maximu...	Safety Set	ADVS	VS Excel Mock Shells
L	16.2.8.2.x	Vital Signs Data	Safety Set	ADVS	PP Excel Mock Shells

1 to 11 of 12 | Page 1 of 2

Challenges, Best Practices & the Road Ahead

CHALLENGES	BEST PRACTICES	ROAD AHEAD
Non-standardized SAP input	Analyzing patterns, and improving efficiency	Enhancing SAP data quality by standardizing input
Building the repository	Construct and update the repository with completed studies	Maintain an up-to-date repository by constructing it with completed studies and regularly updating its contents.
Constructing a training dataset from scratch	Breaking down the elements and constructing organized information	Keep enhancing the dataset
Building custom models	Mastering the intricacies of the human brain to teach machines to think alike	Refine the models through iterative process

Significant Takeaway



01 Accelerate the TLF generation process

02 Optimally utilize time and resources

03 Establish and manage a consistent repository throughout the entire process

Acknowledgement

- Debjit Biswas
- Priya Venkatesan Iyer
- Jeba Kumar
- Periasamy Karuppannan
- Aishwarya L R
- Abhinav Srivastava
- Francis Kevin Raj

The Global Healthcare Data Science Community

Contact Channels
UK +44 1843 609600
office@phuse.global
phuse.global

Social Media
@phusetwitta
/phusebook
/phusetube
/company/phuse

An Overview of Natural Language Processing and Deep Learning

- Deep Learning is a branch of Artificial Intelligence (AI) that focuses on training computers to acquire knowledge from extensive datasets in order to carry out specific tasks. By utilizing neural networks, Deep Learning emulates the cognitive capabilities of the human brain, enabling the identification of meaningful patterns for making informed decisions.

- NLP, a branch of linguistics and AI, is dedicated to comprehending all aspects of human language. The objective of NLP tasks extends beyond the mere comprehension of individual words, aiming to grasp the contextual meaning behind them.