

We claim:

1. A method of extracting information from a technical manual, the method being implemented on a computer system comprising one or more processors, the method comprising:

extracting, by a computer system, selected keywords from the technical manual;

appending, by the computer system, the selected keywords to a table to provide master data;

receiving, by the computer system, a portion of the master data and a first table of labels corresponding to the portion of the master data;

applying, by the computer system, a Machine Learning model or a Natural Language Processing (NLP) model to generate a second table of labels corresponding to an entire master data;

processing, by the computer system, labels in the second table of labels to correspond to the master data; and

outputting, by the computer system, an updated second table of labels to match with the master data to extract information from the technical manual so as to enable a user to review or to use, or both, the information contained in the technical manual by reading labels of the updated second table of labels without reading or searching in the entire technical manual,

wherein the labels in the updated second table of labels are organized according to a maintenance checklist or a suggestion for parameter adjustment, and a data-size of the technical manual is at least five times greater than a data-size of the updated second table of labels.

2. The method of claim 1, wherein extracting keywords from the technical manual comprises using a regular or a rational expression-based extraction procedure.

3. The method of claim 1, wherein the master data comprises Master Inspection Characteristics (MICs) information from the technical manual.

4. The method of claim 1, wherein appending the selected keywords to the table to provide the master data comprises appending the selected keywords to a spreadsheet file.

5. The method of claim 1, wherein, prior to receiving, by the computer system, the portion of the master data and the first table of labels corresponding to the portion of the master data, manually labelling the portion of the master data to generate the first table of labels.
6. The method of claim 1, further comprising inputting, in the computer system, a result of a chi-squared test to the machine learning model to obtain the second table of labels.
7. The method of claim 1, wherein the Machine Learning model comprises a Linear Support Vector Machine (SVM) Classifier.
8. The method of claim 1, wherein the master data comprises master inspection characteristics (MICs) information.
9. The method of claim 1, further comprising, prior to appending, by the computer system, the selected keywords to a table to provide the master data, creating domain-based lexicons and appending, by the computer system, the selected keywords to the table to provide the master data based on the domain-based lexicons.
10. The method of claim 1, wherein the data-size of the technical manual is between 10MB and 100MB, and the data-size of the second table of labels is between 1MB and 2MB.
11. The method of claim 1, wherein the technical manual is a manual of an aircraft engine, a manual of an aircraft system, or both.
12. The method of claim 1, further comprising self-tuning, by the computer system, the machine learning model based on the updated second table of labels corresponding to the master data.
13. The method of claim 1, further comprising, prior to extracting the selected keywords, converting the technical manual from a first file type to a second file type.
14. The method of claim 13, wherein the first file type is a Standard Generalized Markup (SGM) or a Portable Document Format (PDF), and the second file type is an Extensible Markup Language (XML).

15. The method of claim 1, further comprising performing, by the computer system, a text pre-processing of the master data to format the master data to provide pre-processed data, wherein performing, by the computer system, the text pre-processing of the master data comprises performing the text pre-processing of the master data using PANDAS to provide PANDAS pre-processed data.

16. The method of claim 15, wherein performing, by the computer system, the text pre-processing of the master data comprises removing null elements and stop words in the PANDAS pre-processed data to provide cleaned data.

17. The method of claim 16, wherein performing, by the computer system, the text pre-processing of the master data comprises categorizing the cleaned data to provide the pre-processed data.

18. The method of claim 1, wherein applying, by the computer system, the Machine Learning model comprises applying, by the computer system, the Machine Learning model to pre-processed data to generate the second table of labels corresponding to the entire master data comprises converting the pre-processed data into Term Frequency – Inverse Document Frequency (TF-IDF) features to obtain a TF-IDF data.

19. The method of claim 18, wherein applying, by the computer system, the Machine Learning model comprises applying, by the computer system, the Machine Learning model to the pre-processed data to generate the second table of labels corresponding to the entire master data comprises applying a chi-squared test to determine a correlation in the TF-IDF data.

20. A non-transitory computer-readable medium storing a computer-executable code that, when executed by a computer system, causes the computer system to perform a method of extracting information from a technical manual, the method comprising:

extracting, by a computer system, selected keywords from the technical manual;

appending, by the computer system, the selected keywords to a table to provide master data;

receiving, by the computer system, a portion of the master data and a first table of labels corresponding to the portion of the master data;

performing, by the computer system, a text pre-processing of the master data to format the master data to provide pre-processed data;

applying, by the computer system, a Machine Learning model or a Natural Language Processing (NLP) model to the pre-processed data to generate a second table of labels corresponding to an entire master data; and

processing, by the computer system, labels in the second table of labels to correspond to the master data; and

outputting, by the computer system, an updated second table of labels to match with the master data to extract information from the technical manual so as to enable a user to review or to use, or both, the information contained in the technical manual by reading labels of the updated second table of labels without reading or searching in the entire technical manual,

wherein the labels in the updated second table of labels are organized according to a maintenance checklist or a suggestion for parameter adjustment, and a data-size of the technical manual is at least five times greater than a data-size of the updated second table of labels.