

APACHE POI WORD - OVERVIEW

Many a time, a software application is required to generate reference documents in Microsoft Word file format. Sometimes, an application is even expected to receive Word files as input data.

Any Java programmer who wants to produce MS-Office files as output must use a predefined and read-only API to do so.

What is Apache POI?

Apache POI is a popular API that allows programmers to create, modify, and display MS-Office files using Java programs. It is an open source library developed and distributed by Apache Software Foundation to design or modify MS-Office files using Java program. It contains classes and methods to decode the user input data or a file into MS-Office documents.

Components of Apache POI

Apache POI contains classes and methods to work on all OLE2 Compound documents of MS-Office. The list of components of this API is given below:

- **POIFS** *PoorObfuscationImplementationFileSystem* : This component is the basic factor of all other POI elements. It is used to read different files explicitly.
- **HSSF** *HorribleSpreadSheetFormat* : It is used to read and write .xls format of MS-Excel files.
- **XSSF** *XMLSpreadSheetFormat* : It is used for .xlsx file format of MS-Excel.
- **HPSF** *HorriblePropertySetFormat* : It is used to extract property sets of the MS-Office files.
- **HWPF** *HorribleWordProcessorFormat* : It is used to read and write .doc extension files of MS-Word.
- **XWPF** *XMLWordProcessorFormat* : It is used to read and write **.docx** extension files of MS-Word.
- **HSLF** *HorribleSlideLayoutFormat* : It is used to read, create, and edit PowerPoint presentations.
- **HDGF** *HorribleDiaGramFormat* : It contains classes and methods for MS-Visio binary files.
- **HPBF** *HorriblePuBlisherFormat* : It is used to read and write MS-Publisher files.

This tutorial guides you through the process of working on MS-Word files using Java. Therefore the discussion is confined to **HWPF** and **XWPF** components.

Loading [MathJax]/jax/output/HTML-CSS/jax.js