

3rd party parser plugins

List of 3rd party parser plugins

These are 3rd party parser plugins which cannot be included due to licensing incompatibility. To install a plugin, download it according to instructions below and drop the jar(s) on your classpath. Tika will auto detect the plugin.

Microsoft TNEF / LZFU

This is a MS compression format used for compressed RTF, email attachments (like WINMAIL.DAT) and more. The parser is available from a github fork of the [JTNEF project](#).

(Tika 0.10 includes a TNEF parser as standard now, which may be sufficient)

Install instructions:

- `git clone http://github.com/jukka/jtnef.git jtnef`
- `cd jtnef`
- `mvn package`
- `cp target/jtnef-*.jar $SOMEWHERE_ON_CLASS_PATH`

Microsoft Project

This parser extracts metadata and content from Microsoft Project (MPP and MPX) files

It builds on top of [MPXJ](#), which is available under the LGPL

Installation instructions:

- `git clone git://git.code.sf.net/p/mpxj/mpxj`
- `cd !mpxj`
- `mvn package`
- `cp target/mpxj-*SNAPSHOT.jar $SOMEWHERE_ON_CLASS_PATH`
- `git clone http://github.com/Gagravarr/MPXJ-Tika`
- `cd !MPXJ-Tika`
- `mvn package`
- `cp target/mpxj-tika-*SNAPSHOT.jar $SOMEWHERE_ON_CLASS_PATH`

Ogg Vorbis and FLAC

This parser extracts metadata from Ogg Vorbis and FLAC audio files.

The library and parser are available under the Apache License, so this is now included as part of Tika.

Your plugin

<Your description here>

Install instructions:

- <Your instructions here>