

## Technical MD Extractor

## Technical MD Extractor

The Technical MD Extractor plug-in represents the family of utilities that extracts the technical properties (such as size, encoding, compression) of a file. Each extractor is responsible for handling a specific file format. However, the sharing of generic extractors is possible.

The extractor plug-in exposes all the properties it can extract, and Rosetta uses this information to extract property values of a given file. The system saves the property values to the DNX and updates the `significantProperties` section, which holds standard technical properties (such as from exif and XDP). It updates also the `fileTechnicalMetadataExtraction` section, which holds information on the metadata extraction run itself (for example, agent and plug-in names, errors that occurred during the run, etc.)

### Note

- The extractor must be associated with a format in the format library in order to run.
- The plug-in is executed by the **TechMD extraction** and **Add Representation TechMD extraction** tasks when the task parameter **Technical MD extraction only** is disabled (cleared).

## Plug-in Parameters

Depending on the specific implementation, a plug-in may or may not require parameters. If the plug-in does require parameters, they are populated during installation of a new instance of the plug-in.

## Usage

The association of an MD extractor to a format is done at the format level (format library). For an MD extractor to be available at the format level, it should be assigned to the same Classification Group that the format is assigned to.

Active	Plug-In Type Name	Plug-In Name	Description	Contact Person	Plug-In Version	View	Edit	Delete
1	MDExtractorPlugin	AIFF-hul-1.10	AIFFHULMDExtractorPlugin	Exlibris Ltd	3.0	View	Edit	Delete
2	MDExtractorPlugin	AIFF-hul-1.17	AIFFHULMDExtractorPlugin	Exlibris Ltd	5.0	View	Edit	Delete
3	MDExtractorPlugin	AIFF-hul	AIFFHULMDExtractorPlugin	Exlibris Ltd	2.2	View	Edit	Delete
4	MDExtractorPlugin	nz.govt.natlib.adapter.arc.ArcAdapter	ArcExtractorPlugin	National Library Ne...	4.0	View	Edit	Delete
5	MDExtractorPlugin	ASCI-hul	ASCIHULMDExtractorPlugin	Exlibris Ltd	2.2	View	Edit	Delete
6	MDExtractorPlugin	ASCI-hul-1.17	ASCIHULMDExtractorPlugin	Exlibris Ltd	5.0	View	Edit	Delete
7	MDExtractorPlugin	ASCI-hul-1.10	ASCIHULMDExtractorPlugin	Exlibris Ltd	3.0	View	Edit	Delete
8	MDExtractorPlugin	nz.govt.natlib.adapter.bmp.BitmapAdap...	BitmapExtractorPlugin	Exlibris Ltd	3.0	View	Edit	Delete
9	MDExtractorPlugin	BYTESTREAM-hul	BYTESTREAMHULMDExtractor	Exlibris Ltd	2.2	View	Edit	Delete
10	MDExtractorPlugin	BYTESTREAM-hul-1.17	BYTESTREAMHULMDExtractor	Exlibris Ltd	5.0	View	Edit	Delete
11	MDExtractorPlugin	BYTESTREAM-hul-1.10	BYTESTREAMHULMDExtractor	Exlibris Ltd	3.0	View	Edit	Delete

### MD Extractor Plug-In Management

After the MD extractor is assigned to a classification, it is listed in the Related MD extractors folder of that classification group.

Home / Preservation: Classification Groups / Details

<b>Name</b>	Audio (AES)	<b>Created on</b>	06-10-2009	<b>Updated on</b>	03-12-2009
<b>Description</b>	For all uncompr...	<b>Created by</b>	PMcKinney	<b>Updated by</b>	Ex Libris

General Details | **Related MD Extractors** | Related Format Validators | Related Properties | Related Formats | History

MD Extraction Tool	Description
1 AIFF-hul	AIFFHULMDExtractorPlugin
2 AIFF-hul-1.10	AIFFHULMDExtractorPlugin
3 WAVE-hul	WaveHULMDExtractorPlugin
4 WAVE-hul-1.10	WaveHULMDExtractorPlugin
5 nz.govt.natlib.adapter.wav.WaveAdapter	WaveExtractorPlugin

Switch to Local Cancel

### Metadata Extractor Tools

The same list of MD extractors is also available in the MD extractor drop-down at the format level for all formats belonging to the same classification.

Home / Preservation: Formats / Details

<b>Name</b>	fmt/14	<b>Created on</b>	12-03-2009	<b>Updated on</b>	12-03-2009
<b>Description</b>	Portable Docume...	<b>Created by</b>	Ex Libris	<b>Updated by</b>	Ex Libris
<b>Classification</b>	PDF				

General Details | **Related Applications** | Risk Identifiers | Sustainability Factors | Notes | History

\* Name: fmt/14

\* Description: Portable Document Format

Version: 1

Other Names: PDF(1.0)

File Extensions: pdf

Mime Type: application/pdf

Registry Type: PRONOM

\* Registry ID: fmt/14

Registry Second: 613

Classification: PDF

MD Extractor: PDF-hul-1.24 (not listed)

Format Validator: PDF-hul (PDFHULMDExtractorPlugin)

Obsolete: PDF-hul-1.10 (PDFHULMDExtractorPlugin)

End Support Date: nz.govt.natlib.adapter.pdf.PDFAdapter (PDFExtractorPlugin)

Developer: nz.govt.natlib.adapter.pdfbox.PDFBoxAdapter (PDFBoxExtractorPlugin)

Supported By: PDF-hul-1.24 (not listed)

### Metadata Extractor Drop-down Menu

## Implementations

Rosetta includes the following implementations of the Technical MD Extractor plug-in:

- JHOVE
  - BYTESTREAM-hul
  - ASCII-hul
  - AIFF-hul
  - HTML-hul
  - JPEG2000-hul
  - PDF-hul
  - JPEG-hul
  - GIF-hul
  - TIFF-hul
  - UTF8-hul
  - WAVE-hul
  - XML-hul
- NLNZ extraction tool
  - nz.govt.natlib.adapter.flac.FlacAdapter
  - nz.govt.natlib.adapter.bmp.BitmapAdapter
  - nz.govt.natlib.adapter.mp3.MP3Adapter
  - nz.govt.natlib.adapter.arc.ArcAdapter
  - nz.govt.natlib.adapter.wav.WaveAdapter
  - nz.govt.natlib.adapter.pdf.PDFAdapter
  - nz.govt.natlib.adapter.jpg.JpgAdapter
  - nz.govt.natlib.adapter.openoffice.OpenOfficeAdapter
  - nz.govt.natlib.adapter.pdfbox.PDFBoxAdapter
  - nz.govt.natlib.adapter.any.DefaultAdapter
  - nz.govt.natlib.adapter.works.DocAdapter
  - nz.govt.natlib.adapter.excel.ExcelAdapter
  - nz.govt.natlib.adapter.gif.GIFAdapter
  - nz.govt.natlib.adapter.html.HTMLAdapter
  - nz.govt.natlib.adapter.powerpoint.PowerPointAdapter
  - nz.govt.natlib.adapter.tiff.TIFFAdapter
  - nz.govt.natlib.adapter.word.WordAdapter
  - nz.govt.natlib.adapter.wordperfect.WPAdapter
  - nz.govt.natlib.adapter.xml.XMLAdapter3
  - nz.govt.natlib.adapter.xml.XMLAdapter