HTML Metadata Extractor Introduction

- [Features](#)

The HTML Metadata Extractor component allows the extraction of metadata from HTML pages. These pages would most likely come from Aspider Web Crawler. The metadata extracted will be store inside the document processed.

Features

Some of the features of the Publish to HTML Metadata Extractor component include:

- Extract metadata from documents with HTML content.
- Support for Jsoup selectors to extract the metadata