

Tika protocol handler

The Tika protocol handler allows to use the [Apache Tika](#) framework to parse arbitrary documents.

Options

none

Example

To access the different attribute values the [KeyValueToTuple](#) operator can be used to transform the required attributes into a relational tuple.

PQL

Example

```
input = ACCESS({source='source',
                wrapper='GenericPush',
                transport='directory',
                protocol='tika',
                dataHandler='KeyValueObject',
                options=[['directory', '/var/log']]
})

out = KEYVALUETOTUPLE({
    schema=[['content', 'String']],
    type='Document',
    keepinput='false'
},
input
)
```