

SourceTypeAction (content aware pipelines)

1. Introduction

SourceTypeAction assigns a "type" (a string) to an XML file. This is done based on information occurring in the header of the XML file, up to the document (root) element. This type is then returned to the sitemap as a variable with the name 'sourcetype'. If no matching sourcetype could be found, null is returned and thus the contents of the action element will not be executed.

SourceTypeAction works by pull-parsing the document and collecting information such as the Public Id, the processing instructions, the document element local name and namespace, and the `xsi:schemaLocation` and `xsi:noNamespaceSchemaLocation` attributes. This information is then compared with the rules described in the configuration of the SourceTypeAction.

2. Configuration

The action should be declared and configured in the `map:actions` section of the sitemap. Example:

```
<map:action name="sourcetype" src="org.apache.cocoon.acting.sourcetype.SourceTypeAction">
  <sourcetype name="documentv11">
    <document-declaration public-id="-//APACHE//DTD Documentation V1.1//EN"/>
  </sourcetype>
  <sourcetype name="foo">
    <document-element local-name="foo">
  </sourcetype>
</map:action>
```

Each `sourcetype`-tag declares a source type. Inside the `sourcetype`-tag a number of rules can be defined, described below. The sourcetypes will be checked in the same order as they are defined in the configuration, the first sourcetype of which all rules match will be used.

These are the available rules:

document-declaration

This rule checks the public ID. It takes one attribute `public-id`.

document-element

This rule checks the local name and/or namespace of the document element. These are specified with the attributes `local-name` and `namespace`. At least one of these two is required.

processing-instruction

This rule checks a processing instruction. It can take two attributes: `target` and `data`. The `target` attribute is always required, the `data` attribute is optional.

w3c-xml-schema

This rule checks the value of the `xsi:schemaLocation` and `xsi:noNamespaceSchemaLocation` attributes on the document element. These are specified with the attributes `schema-location` and `no-namespace-schema-location`.

3. Usage

The source of which the sourcetype must be defined must be specified using the 'src' attribute on the `map:act` element.

```
<map:act type="sourcetype" src="{1}">
  ...
</map:act>
```