Source

Cocoon often uses URLs to access data, such as in the `src` attribute of `<map:generate>` and `<map:transform>`. These URLs are named "sources", because they go far beyond standard JDK URLs...

What is a source ?

A source is a means to access data, and makes no assumption on what kind of data it is. It provides the low-level plumbing to access the data. That's why we have FileSource, URLSource, BlobSource (for SQL blobs), SlideSource (for webdav), etc. There is absolutely no assumption that data accessed using a Source is XML, nor that it can be converted to XML (e.g. gif images).

Sources were added to Cocoon, and later to Avalon because the URL protocol extension mechanism provided by the JDK is very limited, as it requires some parameters to be set at the JVM level and only allows JVM-wide protocol definitions, which is obviously bad for application servers.

Cocoon comes with a number of source implementations which define the Protocols you're used to when writing a sitemap.

Resolving a source

TODO explain SourceResolver, SourceFactory and URLFactory

Obtaining XML from a Source

We said above that a Source can provide access to any kind of data, and not only XML. However, Cocoon being an XML framework, we often need to obtain XML from a source.

To obtain XML from a source, you will use the `toSAX()` method of the `SourceResolver`. This method tries to produce XML by following these steps :

- if the source is XMLizable (i.e. can natively produce XML), then use the source's `toSAX` method.
- if the source can determine it's content type using `getMimeType()`, try to find an applicable XMLizer. An XMLizer is an Avalon component that converts data of a given mime-type to XML. For example, sources of type "text/html" can be converted to XML using JTidy.
- finally, consider the data is XML text and feed an XML parser with this data.

Sources vs Generators

A question that comes often when people want to add a new way of obtaining data to Cocoon is "should I write a Source or a Generator ?". The answer is "SourceVsGenerator"...