ADAMS

Advanced **D**ata mining **A**nd **M**achine learning **S**ystem

Module: adams-xml

Peter Reutemann

January 10, 2024

©2019

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Except where otherwise noted, this work is licensed under
http://creativecommons.org/licenses/by-sa/4.0/

Contents

1 Flow **5**

2 Tools **7**
2.1 Pretty print XML 7

Bibliography **9**

Chapter 1

Flow

ADAMS offers some basic handling for XML[2], XSLT[3] and XPath[4]. The following sources are available:

- *NewDOMDocument* – creates an empty DOM document.

The following transformers are available:

- *AddDOMAttribute* – adds an attribute and its value to the node passing through.
- *AddDOMNode* – appends a child new to the node passing through.
- *XMLFileReader* – reads the specific XML file and forwards a DOM document.
- *XPath* – applies an XPath expression to the incoming DOM document.
- *XSLT* – applies an XSLT stylesheet to the incoming DOM document.

The following sinks are available:

- *DOMDisplay* – simple tree view of a DOM node.
- *XMLFileWriter* – writes the DOM document to disk.

The following conversions are available:

- *DOMToString* – turns the DOM object into an XML string.
- *DOMToProperties* – flattens the DOM object into a Java Properties object (key-value pairs).
- *DOMNodeToString* – turns the DOM node into an XML string.
- *DOMNodeListToArray* – turns the list of DOM nodes into an array.
- *PrettyPrintXML* – turns an XML string into a pretty-printed XML string.
- *XMLToDOM* – parses the XML string and generates a DOM object.

Chapter 2

Tools

2.1 Pretty print XML

In order to preserve space, XML is often optimized and removes any unnecessary whitespaces. However, for a human to inspect such data, it is much more useful to have it properly indented, aka *pretty printed*. Figure 2.1 shows a screenshot of the *Pretty print XML* tool that allows you to turn XML into a more human-readable format.

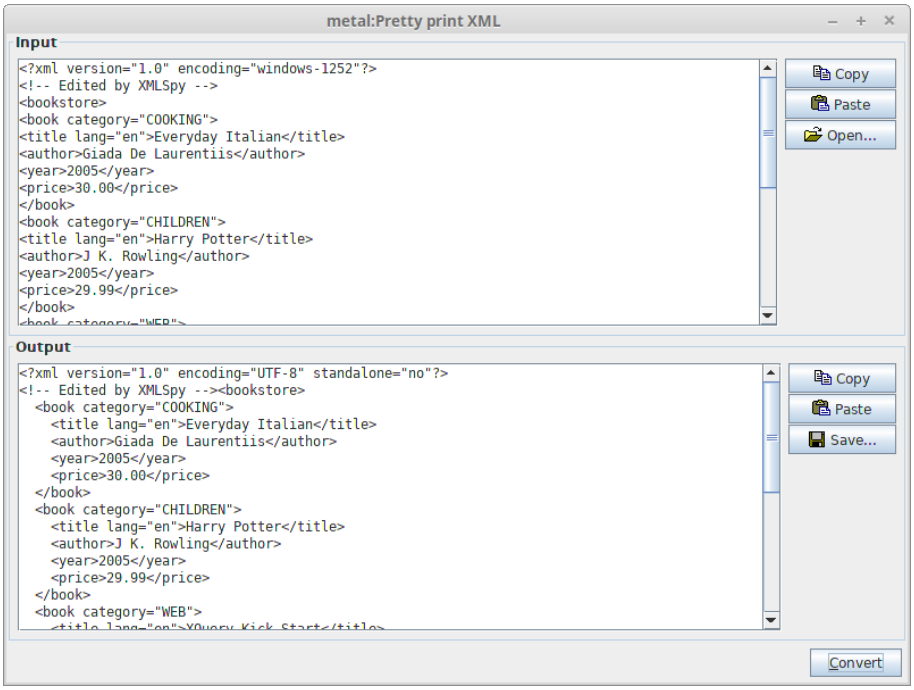

Figure 2.1: Pretty print XML

Bibliography

[1] *ADAMS* – Advanced Data mining and Machine learning System
`https://adams.cms.waikato.ac.nz/`

[2] *XML* – Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.
`https://en.wikipedia.org/wiki/XML`

[3] *XSLT* – XSLT (eXtensible Stylesheet Language Transformations) is a language for transforming XML documents into other XML documents or other formats.
`https://en.wikipedia.org/wiki/XSLT`

[4] *XPath* – XPath (XML Path Language) is a query language for selecting nodes from an XML document.
`https://en.wikipedia.org/wiki/XPath`