


This repository | Search

Pull requests Issues Marketplace Gist

 mckamey / jsonml

Watch 9 Star 123 Fork 17

<> Code ⓘ Issues 3 🏷️ Pull requests 2 📁 Projects 0 📖 Wiki Insights ▾

Add support for Processing instructions #14

Edit

Open Holzhaus wants to merge 5 commits into mckamey:master from RUB-NDS:processing-instructions

🗨️ Conversation 0 🔑 Commits 5 📄 Files changed 3

+68 -4



Holzhaus commented an hour ago • edited



This adds support for [Processing Instructions](#). These are defined in the XML 1.0 Specification as follows:

```
[16] PI          ::= '<?' PITarget (S (Char* - (Char* '?>' Char*))? '>'  
[17] PITarget    ::= Name - (('X' | 'x') ('M' | 'm') ('L' | 'l'))
```

Rationale

The [JsonML website](#) states:

Processing instruction nodes are inherently ignored in JsonML. There does not seem to be a useful analogy which is consistent with the intent of JsonML.

But to provide a truly lossless way to convert arbitrary XML to JsonML and back, IMHO they *should be preserved*:

- As per [Section 2.6 of the XML 1.0 Specification](#), passing PIs through to an application is a "MUST" requirement.
- Accordingly, PIs are also preserved during [XML Canonicalization \(C14N\)](#), which is a way to normalize the physical appearance of an XML document while retaining its logic. The retention of PIs indicates that they are not insignificant (by contrast, comment support is optional and things like DTD or XML Declaration are lost during C14N).

JsonML grammar modifications

Thus, I added support for Processing Instructions:

```
element  
= '[' tag-name ',' attributes ',' element-list '']'  
| '[' tag-name ',' attributes ']'  
| '[' tag-name ',' element-list ']'  
| '[' tag-name ']'  
| '[' pi-target ',' string ']'  
| string  
;  
pi-target  
= "?" + string
```

Processing Instructions can't be mistaken for elements since `#x3F` (= Question Mark [?]) is not a valid `NameStartChar` (which means that a `tag-name` can't begin with a question mark):

```
[2] Char      ::= #x9 | #xA | #xD | [#x20-#xD7FF] | [#xE000-#xFFFF] | [#x10000-#x1  
[3] S         ::= (#x20 | #x9 | #xD | #xA)+  
[4] NameStartChar ::= ":" | [A-Z] | "_" | [a-z] | [#xC0-#xD6] | [#xD8-#xF6] | [#xF8-#x  
[4a] NameChar  ::= NameStartChar | "-" | "." | [0-9] | #xB7 | [#x0300-#x036F] | [#x  
[5] Name       ::= 'NameStartChar (NameChar)*'  
[40] STag      ::= '<' Name (S Attribute)* S? '>'
```

One subtlety is that when a document contains one or more Processing Instructions as top-level constructs (besides the `documentElement`), then the resulting JsonML text will have a top level element with an empty string as `tag-name` and an `element-list` consisting of the PIs and the `documentElement` in the correct

Reviewers

No reviews

Assignees

No one assigned

Labels

None yet

Projects

None yet

Milestone

No milestone

Notifications

🔊 Unsubscribe

You're receiving notifications because you authored the thread.

1 participant



☑️ Allow edits from maintainers.

[Learn more](#)

order.
This behaviour is consistent with the [handling of fragments in JsonML](#).
Please have a look at the examples below and let me know what you think. Thanks in advance!

Examples

Here's a simple example:

```
<foo><?a-single-pi with data?></foo>
```

The above XML document represented as JsonML:







```
[
  "foo",
  [
    "?a-single-pi",
    "with data"
  ]
]
```

Here's a more complex example with has Processing Instructions at the document level:

```
<?some-pi and its data?>
<foo>
  <?another-pi with data?>
</foo>
<?third-pi?>
```

The above file will map to:

```
[
  "",
  [
    "?some-pi",
    "and its data"
  ],
  "\n",
  [
    "foo",
    "\n  ",
    [
      "?another-pi",
      "with data"
    ],
    "\n"
  ],
  "\n",
  [
    "?third-pi",
    ""
  ]
]
```

-  Holzhaus added some commits 11 days ago
-  Add support for ProcessingInstruction nodes 73e5650
 -  Also use ProcessingInstructions outside the documentElement ... c6a0b50
 -  Avoid empty text node parent if documentElement is the only top-level... 0a75a40
 -  Add unittests for processing instructions edbbc25
 -  Add support for Processing Instructions to jsonml-utils.js aa601ae

Add more commits by pushing to the **processing-instructions** branch on **RUB-NDS/jsonml**.

This branch has no conflicts with the base branch

Only those with [write access](#) to this repository can merge pull requests.



Write Preview

AA B i “ < > ↺ ⋮ ≡ ✓ ↶ @ 📌

Leave a comment


Attach files by dragging & dropping, [selecting them](#), or pasting from the clipboard.

Md

Styling with Markdown is supported

Close pull request

Comment

 **ProTip!** Add comments to specific lines under [Files changed](#).