JWAT v0.9.0 (2012-05-11)

Released 2012-05-11.

- **Highlights**
  - **GZip package**
    - Refactored GZip reader/validator
    - GZip writer
    - Input/Output stream
    - Validation reporting
    - Unit tests completeness
  - **ARC/WARC**
    - More unit tests
    - Unified the reader API
  - **Common package**
    - Unit testing of Stream classes
    - Completed the HeaderLineReader
  - Full list of issues resolved in this release.

**Highlights**

**GZip package**

Refactored GZip reader/validator

The reader/validator was completely rewritten since the original one was badly implemented.

**GZip writer**

A GZip writer was implemented for completeness and also to test the reader fully.

**Input/Output stream**

The reader/writer now has support for streams directly.

**Validation reporting**

Warnings/Errors are now reported in the same manor as the ARC/WARC packages.

**Unit tests completeness**

Added more unit tests so also all GZip code is test covered.

**ARC/WARC**

More unit tests

Added better test coverage but still not completely covered.

**Unified the reader API**

The ARC/WARC factories/readers should now have a similar API.

**Common package**

Unit testing of Stream classes

The stream classes were not completely unit tested in the previous release, but they should be now.

**Completed the HeaderLineReader**

The HeaderLineReader now includes validation of RAW, US-ASCII, ISO-8859-1, UTF-8, Quoted-Strings and Encoded Words.

**Full list of issues resolved in this release.**

<table>
<thead>
<tr>
<th>T</th>
<th>Key</th>
<th>P</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>JWAT-19</td>
<td></td>
<td>Complete WARC headerline reader support for utf-8, quoted-strings and encoded words.</td>
</tr>
<tr>
<td>+</td>
<td>JWAT-20</td>
<td></td>
<td>Unit tests and coverage for utf-8, quoted-strings and encoded words support.</td>
</tr>
<tr>
<td>+</td>
<td>JWAT-13</td>
<td></td>
<td>Unit tests and coverage of GZip code</td>
</tr>
<tr>
<td>+</td>
<td>JWAT-11</td>
<td></td>
<td>Refactor HttpResponse parser code</td>
</tr>
<tr>
<td>#</td>
<td>JWAT</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>JWAT-9</td>
<td>Refactor GZip code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-41</td>
<td>ArcReader/Record does not seem to return errors/warnings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-42</td>
<td>Refactor and unify gzip, arc and warc error/warning handling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-40</td>
<td>GZip report errors/warnings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-39</td>
<td>HttpResponse uses mark/reset which have been disabled in the refactored streams.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-29</td>
<td>Unit tests and coverage of streams in streams (validate consumption)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-27</td>
<td>Unit tests and coverage of HttpResponse in a Payload</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-22</td>
<td>Unify ARC and WARC Reader/Factory Api</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JWAT-10</td>
<td>Implement GZip writer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

13 issues