

Infrastructure - Bug #4225

Do not parse any RDF object as an ORE object

2014-01-14 17:22 - Ben Leinfelder

| | |
|--|---|
| Status: Closed | Start date: |
| Priority: Normal | Due date: |
| Assignee: Skye Roseboom | % Done: 100% |
| Category: d1_indexer | Estimated time: 0.00 hour |
| Target version: | Story Points: |
| Milestone: CCI-1.3 | |
| Product Version: * | |
| Description | |
| <p>I've been working with the CN indexing libraries and noticed that the ResourceMapSubprocessor is "handling" both these formatIds: http://www.openarchives.org/ore/terms http://www.w3.org/TR/rdf-syntax-grammar</p> <p>But I believe it should only be handling the first one since many assumptions are made about the object and the triples it contains.</p> <p>Also, I am trying to make a more generalized RDF/XML processor that would allow us to pick off concepts for indexing and when two subprocessors claim to handle the same formatId, only one of them can consume the stream - the next one gets an error. We could consider reconstructing the stream to support multiple parsers for the same document format to support this as well.</p> | |
| Related issues: | |
| Related to Infrastructure - Story #4461: CCI 1.2.6 Feature Release | Closed 2014-03-19 2014-05-10 |

History

#1 - 2014-03-04 15:14 - Skye Roseboom

- Status changed from New to In Progress

#2 - 2014-03-04 15:21 - Skye Roseboom

- Tracker changed from Story to Bug

- Status changed from In Progress to Testing

Moved configuration correction to build out on trunk/unstable for testing.

#3 - 2014-05-01 14:56 - Skye Roseboom

- Status changed from Testing to Closed

Update to configuration of the ResourceMapSubProcessor has been updated to only handle '<http://www.openarchives.org/ore/terms>'