

Ingest XML files fails in some servers

- **Article Type:** General
 - **Product:** Rosetta
-

Problem Symptoms:

Some xml files are going through correctly in one environment and fail format validation in the other.

The error we are seeing in the logs is this:

```
2014-01-03 07:36:17,622 WARN [com.exlibris.dps.repository.plugin.mdExtractor.AbstractJhoveMDExtractorPlugin]
(SIP_PROCESSING_QUEUE Queue Job Receiver 315) [] jhove message: schema_reference.4: Failed to read schema
document 'http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd', because 1) could not find the document; 2)
the document could not be read; 3) the root element of the document is not <xsd:schema>.
```

- Both environment, the file validation uses same JHOVE, XML-hul1.4
- Both environment use the same plugin XML-hul for fmt/101(not 1_10)

Cause:

Firewall issues are stopping JHOVE using a validation xsd

Resolution:

run from the server:

```
>wget http://www.loc.gov/MARC21/slim http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd
```

If you fail to connect, JHOVE will fail identifying the xml format.

- **Article last edited:** 2/17/2014