

Ninestars QA Suite

The Ninestars QA Suite consists of two tools for checking the newspaper scans for The State and University Library in Aarhus, Denmark.

One tool is designed to be run on a single JPEG 2000 file, and will report any deviations from the specification of the file. The other tool is designed to be run on a directory structure containing the output of an entire batch, including data and metadata, and will report any deviations from the specifications.

What is validated

The following things are validated:

- All jp2 files are validated to conform to the specifications
- All md5sums are checked to be correct
- All directories are checked to have the correct names and contain the correct files (this is a work in progress)
- All files are checked to have the correct names (this is a work in progress)
- The [AvisID] and [date] are checked to conform to the data in the database (this is a work in progress)
- All XML files are checked to conform to the correct XML Schema
- Automatic check of contents in XML metadata as specified in Appendix 2C, 2J and 2K (this is a work in progress)

The following things are planned for a later release of the tool:

- Completion of the directory file name validations
- Completion of the [AvisID] and [date] validations
- Completion of automatic check of XML metadata files as specified in Appendix 2C, 2J and 2K
- Automatic check of contents in XML metadata files, as specified in Appendix 2D and 2E

Prerequisites

The QA Suite is a work in progress, and the list of prerequisites can not yet be fully described.

Prerequisites will include:

- A linux system. The tool has been tested on [CentOS](#) release 6.4, but most linux versions should work, provided the other prerequisites can be met.

- [Java SE 7](#)
- [Python](#) version 2.7 or above or 3.2 or above

This list may be extended.

Running

The file tool

Start the file tool with

```
bin/qafile.sh <path-to-file>
```

<path-to-file> is the path to the file to check.

The batch tool

Start the batch tool with

```
bin/qabatch.sh <path-to-batch> <sql-connection-string>
```

<path-to-batch> is the path to the directory containing the batch to check.
Example: “/var/spool/batch/B400022028241-RT1”

<sql-connection-string> is the connection string used to connect to the database with the information about the batch. Example: “jdbc:postgresql://dbhost/mfpak?user=mfpak&password=mfpass”

The database contains information about the data in a batch, and a reference to documentation about the database will be inserted, when it is ready.

Output

Both tools will output an XML file to stdout, containing the result of the validation (success or failure) and a list of errors, if any.

The schema for the XML file can be found here [Schema](#). Note that the current version of the tool is a little lax in it’s use of <component>, <type> and <description> values. These values will be reviewed and improved in a later version of the tool.

Examples of output can be found here: [Success](#) and [Failure](#)

The QA Suite is a work in progress, and the schema and examples may be extended as a result of this.

If the tool fails unexpectedly, it will exit with an error code larger than zero, and error messages will be printed to stderr.

The tool also produces a log while executing. By default the log will be output to stderr with a level of `WARNING`. This can be changed by editing the file `conf/logback.xml`. You can configure logging as described in the [Logback manual](#).